ESAIM: COCV 17 (2011) 243–266
DOI: 10.1051/cocv/2010003

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

A DUALITY-BASED APPROACH TO ELLIPTIC CONTROL PROBLEMS IN NON-REFLEXIVE BANACH SPACES *

Christian Clason[1] and Karl Kunisch[1]

Abstract. Convex duality is a powerful framework for solving non-smooth optimal control problems. However, for problems set in non-reflexive Banach spaces such as $\mathrm{L}^1(\Omega)$ or $\mathrm{BV}(\Omega)$, the dual problem is formulated in a space which has difficult measure theoretic structure. The predual problem, on the other hand, can be formulated in a Hilbert space and entails the minimization of a smooth functional with box constraints, for which efficient numerical methods exist. In this work, elliptic control problems with measures and functions of bounded variation as controls are considered. Existence and uniqueness of the corresponding predual problems are discussed, as is the solution of the optimality systems by a semismooth Newton method. Numerical examples illustrate the structural differences in the optimal controls in these Banach spaces, compared to those obtained in corresponding Hilbert space settings.

Mathematics Subject Classification. 49J52, 49J20, 49K20.

Received March 3, 2009. Revised July 2, 2009.
Published online March 24, 2010.

1. Introduction

This work is concerned with the study of the optimal control problem

$$\begin{cases} \min\limits_{u\in\mathcal{X}} \dfrac{1}{2}\left\|y-z\right\|^2_{\mathrm{L}^2(\Omega)} + \alpha\left\|u\right\|_{\mathcal{X}} \\ \text{s.t. } Ay = u \end{cases} \tag{$\mathcal{P}$}$$

where $\Omega \subset \mathbb{R}^n$, $n \in \{2,3\}$, is a simply connected bounded domain with Lipschitz boundary $\partial\Omega$, $\mathcal{X}$ is a non-reflexive Banach space, and $\alpha > 0$ and $z \in \mathrm{L}^2(\Omega)$ are given. Furthermore, A is a linear second order elliptic differential operator, taking appropriate boundary conditions. Throughout, we also assume that:

$$\left\|A\cdot\right\|_{\mathrm{L}^2} \text{ and } \left\|A^*\cdot\right\|_{\mathrm{L}^2} \text{ are equivalent norms on } \mathrm{H}^2(\Omega)\cap\mathrm{H}^1_0(\Omega), \tag{A}$$

where A^* denotes the adjoint of A with respect to the inner product in $\mathrm{L}^2(\Omega)$.

Keywords and phrases. Optimal control, L^1, bounded variation (BV), measures, Fenchel duality, semismooth Newton.

* *This work was supported by the Austrian Science Fund (FWF) under grant SFB F32 (SFB "Mathematical Optimization and Applications in Biomedical Sciences").*

[1] Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36, 8010 Graz, Austria.
christian.clason@uni-graz.at; karl.kunisch@uni-graz.at

Article published by EDP Sciences

© EDP Sciences, SMAI 2010

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

If $\mathcal{X} = \mathrm{L}^1(\Omega)$, this setting applies to optimal control problems where the cost of the control is a linear function of its magnitude (*cf.* [15]); the case $\mathcal{X} = \mathrm{BV}(\Omega)$ corresponds to settings where the cost is proportional to changes in the control. Of particular interest is how the structure of optimal controls in such Banach spaces differs from that of controls obtained in Hilbert spaces such as $\mathrm{L}^2(\Omega)$ or $\mathrm{H}^1_0(\Omega)$. For example, it is known that $\mathrm{L}^1(\Omega)$-type costs promote sparsity, whereas $\mathrm{BV}(\Omega)$-type penalties favor piecewise constant functions (*cf.* [10,11], respectively). Note that for $\mathcal{X} = \mathrm{L}^1(\Omega)$, problem ($\mathcal{P}$) is not well-posed: It need not have a minimizer in $\mathrm{L}^1(\Omega)$, since the conditions of the Dunford-Pettis theorem are not satisfied (boundedness in $\mathrm{L}^1(\Omega)$ is not a sufficient condition for the existence of a weakly converging subsequence). The natural functional-analytic framework for problems of this type is the space of bounded measures (*cf.* Rem. 2.8). In the current paper, we will focus on optimal control problems in the space of measures and in the space of functions of bounded variation, and on their numerical treatment.

For the direct solution of ($\mathcal{P}$), one would need to address the problem of the discretization of measures. We therefore propose an alternative approach that, roughly speaking, consists in interpreting the optimality conditions for problem ($\mathcal{P}$) as an optimality system for the Lagrange multiplier associated with the equality constraint, which by a density argument can then be taken in an appropriate Hilbert space. This can be justified rigorously using Fenchel duality. This, together with the numerical results for simple model problems which highlight the significant difference of the controls in dependence of the chosen norm, constitutes the main contribution of the current paper.

This work is organized as follows: In the remainder of this section, we fix notations and recall some necessary background. In Section 2, we derive predual formulations for the optimal control problem ($\mathcal{P}$) with measures and functions of bounded variation as controls (in Sects. 2.1 and 2.2, respectively), discuss the existence and uniqueness of their solutions, and derive optimality systems. Section 3 is concerned with the solution of the optimality systems by a semismooth Newton method, for which it is necessary to consider a regularization of the problem (Sect. 3.1). We can then show superlinear convergence of the method (Sect. 3.2). Here, we focus first on the case of measures, and discuss the corresponding issues for functions of bounded variation in Section 3.3. Finally, we present numerical examples in Section 4.

1.1. Notations and background

For the reader's convenience, we give here the definitions and results on measure theory, functions of bounded variation, and convex duality relevant to this work. For more details and proofs, we refer to, *e.g.*, [1,3] (our notation follows the latter). In the following, Lebesgue spaces of vector valued functions are denoted by a bold letter corresponding to their scalar equivalent, *e.g.* $\mathbb{L}^2(\Omega) := (\mathrm{L}^2(\Omega))^n$.

1.1.1. *Measure theory*

Let $\mathcal{M}(\Omega)$ denote the vector space of all bounded Borel measures on Ω, that is of all bounded σ-additive set functions $\mu : \mathcal{B}(\Omega) \to \mathbb{R}$ defined on the Borel algebra $\mathcal{B}(\Omega)$ satisfying $\mu(\emptyset) = 0$. The total variation of $\mu \in \mathcal{M}(\Omega)$ is defined for all $B \in \mathcal{B}(\Omega)$ by

$$|\mu|(B) := \sup \left\{ \sum_{i=0}^{\infty} |\mu(B_i)| : \bigcup_{i=0}^{\infty} B_i = B \right\}, \tag{1.1}$$

where the supremum is taken over all partitions of B. Endowed with the norm $\|\mu\|_{\mathcal{M}} = |\mu|(\Omega)$, $\mathcal{M}(\Omega)$ is a Banach space. By the Riesz representation theorem, $\mathcal{M}(\Omega)$ can be isometrically identified with the topological dual of $\mathrm{C}_0(\Omega)$, the space of all continuous functions with compact support in Ω, endowed with the norm $\|v\|_{\mathrm{C}_0} = \sup_{x\in\Omega} |v(x)|_\infty$. This leads to the following equivalent characterization of the norm on $\mathcal{M}(\Omega)$:

$$\|\mu\|_{\mathcal{M}} = \sup_{\substack{\varphi\in \mathrm{C}_0(\Omega),\\ \|\varphi\|_{\mathrm{C}_0}\leq 1}} \int_\Omega \varphi \,\mathrm{d}\mu. \tag{1.2}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

1.1.2. *Functions of bounded variation*

We recall that $\mathrm{BV}(\Omega)$, the space of functions of bounded variation, consists of all $u \in \mathrm{L}^1(\Omega)$ for which the distributional gradient Du belongs to $(\mathcal{M}(\Omega))^n$. Furthermore, the mapping $u \mapsto \|u\|_{\mathrm{BV}}$,

$$\|u\|_{\mathrm{BV}} := \int_\Omega |Du| \, \mathrm{d}x = \sup \left\{ \int_\Omega u \operatorname{div} v \, \mathrm{d}x : \; v \in (\mathrm{C}_0^\infty(\Omega))^n, \; \|v\|_{(\mathrm{C}_0)^n} \leq 1 \right\} \tag{1.3}$$

(which can be infinite) is lower semicontinuous in the topology of $\mathrm{L}^1(\Omega)$, and $u \in \mathrm{L}^1(\Omega)$ is in $\mathrm{BV}(\Omega)$ if and only if $\|u\|_{\mathrm{BV}}$ is finite. (If $v \in \mathrm{H}^1(\Omega)$, then $\|u\|_{\mathrm{BV}} = \int_\Omega |\nabla u| \, \mathrm{d}x$.) Endowed with the norm $\|\cdot\|_{\mathrm{L}^1} + \|\cdot\|_{\mathrm{BV}}$, $\mathrm{BV}(\Omega)$ is a non-reflexive Banach space.

One of the main features of $\mathrm{BV}(\Omega)$ in comparison to Sobolev spaces is that it includes characteristic functions of sufficiently regular sets and piecewise smooth functions. In image reconstruction, $\mathrm{BV}(\Omega)$-regularization is known to preserve edges better than regularization with $\|\nabla u\|_{\mathrm{L}^2}^2$.

1.1.3. *Fenchel duality in convex optimization*

A complete discussion can be found in [6]. Let V and Y be Banach spaces with topological duals V^* and Y^*, respectively, and let $\Lambda : V \to Y$ be a continuous linear operator. Setting $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$, let $\mathcal{F} : V \to \overline{\mathbb{R}}$, $\mathcal{G} : Y \to \overline{\mathbb{R}}$ be convex lower semicontinuous functionals which are not identically equal ∞ and for which there exists a $v_0 \in V$ such that $\mathcal{F}(v_0) < \infty$, $\mathcal{G}(\Lambda v_0) < \infty$, and $\mathcal{G}$ is continuous at Λv_0. Let $\mathcal{F}^* : V^* \to \overline{\mathbb{R}}$ denote the Fenchel conjugate of $\mathcal{F}$ defined by

$$\mathcal{F}^*(q) = \sup_{v \in V} \langle q, v \rangle_{V^*,V} - \mathcal{F}(v), \tag{1.4}$$

which we will frequently calculate using the fact that

$$\mathcal{F}^*(q) = \langle q, v \rangle_{V^*,V} - \mathcal{F}(v) \quad \text{ if and only if } \quad q \in \partial \mathcal{F}(v). \tag{1.5}$$

Here, $\partial \mathcal{F}$ denotes the subdifferential of the convex function $\mathcal{F}$, which reduces to the Gâteaux-derivative if it exists.

The Fenchel duality theorem states that under the assumptions given above,

$$\inf_{v \in V} \mathcal{F}(v) + \mathcal{G}(\Lambda v) = \sup_{q \in Y^*} -\mathcal{F}^*(\Lambda^* q) - \mathcal{G}^*(-q) \,, \tag{1.6}$$

holds, and that the right hand side of (1.6) has at least one solution. Furthermore, the equality in (1.6) is attained at (v^*, q^*) if and only if

$$\begin{cases} \Lambda^* q^* \in \partial \mathcal{F}(v^*), \\ -q^* \in \partial \mathcal{G}(\Lambda v^*). \end{cases} \tag{1.7}$$

2. Existence and optimality conditions

This section is concerned with the predual problem of ($\mathcal{P}$), discussing the existence and uniqueness of its solution and deriving the first order optimality system. We first consider the case $\mathcal{X} = \mathcal{M}(\Omega)$, and then treat the case $\mathcal{X} = \mathrm{BV}(\Omega)$ in Section 2.2.

2.1. Controls in $\mathcal{M}(\Omega)$

We consider the optimal control problem

$$\begin{cases} \min\limits_{u \in \mathcal{M}(\Omega)} \dfrac{1}{2} \|y - z\|_{\mathrm{L}^2}^2 + \alpha \|u\|_{\mathcal{M}} \\ \text{s.t. } Ay = u. \end{cases} \tag{$\mathcal{P}_{\mathcal{M}}$}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

First, we have to address the well-posedness of the constraint for measure-valued data. We call $y \in \mathrm{L}^1(\Omega)$ a very weak solution of $Ay = u \in \mathcal{M}(\Omega)$ if

$$\int_\Omega y A^*\varphi \,\mathrm{d}x = \int_\Omega \varphi \,\mathrm{d}u \tag{2.1}$$

holds for all $\varphi \in \mathrm{C}_0(\Omega)$ such that $A^*\varphi \in \mathrm{C}_0(\Omega)$. Then, we have the following result [12], Theorem 9.1:

Proposition 2.1. *For $u \in \mathcal{M}(\Omega)$, the equation $Ay = u$ has a unique weak solution which satisfies $y \in \mathrm{W}_0^{1,p}(\Omega)$ for all $1 \le p < \frac{n}{n-1}$. Furthermore, there exists a constant $C > 0$ such that*

$$\|y\|_{\mathrm{W}_0^{1,p}} \le C \|u\|_{\mathcal{M}} \tag{2.2}$$

holds.

Since $\mathrm{W}_0^{1,p}(\Omega)$ is compactly embedded in $\mathrm{L}^2(\Omega)$, $(\mathcal{P}_\mathcal{M})$ is well-defined, and we can show existence of a minimizer:

Proposition 2.2. *Problem $(\mathcal{P}_\mathcal{M})$ has a unique solution $(y^*, u^*) \in \mathrm{L}^2(\Omega) \times \mathcal{M}(\Omega)$.*

Proof. We consider a minimizing sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{M}(\Omega)$ of $(\mathcal{P}_\mathcal{M})$. Since the pair $(y, u) = (0, 0)$ is feasible, $\frac{1}{2\alpha}\|z\|_{\mathrm{L}^2}^2$ is an upper bound for $\|u_n\|_\mathcal{M}$. We can therefore extract a subsequence converging in the weak topology $\sigma(\mathcal{M}(\Omega), \mathrm{C}_0(\Omega))$ to a $u^* \in \mathcal{M}(\Omega)$.

Setting $y_n := y(u_n) \in \mathrm{W}_0^{1,p}(\Omega)$, *i.e.*, the solution of (2.1) with $\mu = u_n$, we therefore can pass to the limit and obtain (by the density of $\mathrm{C}_0(\Omega)$ in $\mathrm{L}^2(\Omega)$) a $y^* \in \mathrm{W}_0^{1,p}(\Omega) \subset \mathrm{L}^2(\Omega)$ solving (2.1) for u^*. From the weak lower semicontinuity of the norms in $\mathrm{L}^2(\Omega)$ and $\mathcal{M}(\Omega)$, we conclude that (y^*, u^*) is a minimizer of $(\mathcal{P}_\mathcal{M})$.

Finally, uniqueness of a minimizer follows directly from strict convexity of the norms and the assumption on A. □

We set $\mathcal{W} := \mathrm{H}^2(\Omega) \cap \mathrm{H}_0^1(\Omega)$ and consider the problem

$$\begin{cases} \min\limits_{p\in\mathcal{W}} \dfrac{1}{2}\|A^*p + z\|_{\mathrm{L}^2}^2 - \dfrac{1}{2}\|z\|_{\mathrm{L}^2}^2 \\ \text{s.t.} \quad \|p\|_{\mathrm{C}_0} \le \alpha, \end{cases} \tag{$\mathcal{P}_\mathcal{M}^*$}$$

which we will show below to be the predual of $(\mathcal{P}_\mathcal{M})$. Due to the embedding of $\mathcal{W}$ in $\mathrm{C}_0(\Omega)$, the constraint is well-defined, and problem $(\mathcal{P}_\mathcal{M}^*)$ has a unique solution:

Theorem 2.3. *Problem $(\mathcal{P}_\mathcal{M}^*)$ has a unique solution $p^* \in \mathcal{W}$.*

Proof. Let again $\{p_n\}_{n\in\mathbb{N}}$ be a minimizing sequence which is bounded in $\mathcal{W}$ by $2\|z\|_{\mathrm{L}^2}^2$. We can thus extract a subsequence $\{p_{n_k}\}_{k\in\mathbb{N}}$ of feasible functions weakly converging to a $p^* \in \mathcal{W}$, which is again feasible. By the weak lower semicontinuity of the norm, we deduce

$$\liminf_{k\to\infty} \frac{1}{2}\|A^*p_{n_k} + z\|_{\mathrm{L}^2}^2 \ge \frac{1}{2}\|A^*p^* + z\|_{\mathrm{L}^2}^2 . \tag{2.3}$$

Hence, p^* is a minimizer of $(\mathcal{P}_\mathcal{M}^*)$.

By the assumption on A^*, the mapping $p \mapsto \frac{1}{2}\|A^*p + z\|_{\mathrm{L}^2}^2$ is strictly convex, and the minimizer is unique. □

Theorem 2.4. *The dual of $(\mathcal{P}_\mathcal{M}^*)$ is $(\mathcal{P}_\mathcal{M})$, and the solutions $u^* \in \mathcal{M}(\Omega)$ of $(\mathcal{P}_\mathcal{M})$ and $p^* \in \mathcal{W}$ of $(\mathcal{P}_\mathcal{M}^*)$ are related by*

$$\begin{cases} \langle u^*, v\rangle_{\mathcal{W}^*,\mathcal{W}} = \langle A^*p^* + z, A^*v\rangle_{\mathrm{L}^2} , \\ 0 \ge \langle -u^*, p - p^*\rangle_{\mathcal{M},\mathrm{C}_0} , \end{cases} \tag{2.4}$$

for all $v, p \in \mathcal{W}$ with $\|p\|_{\mathrm{C}_0} \le \alpha$.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Proof. We apply Fenchel duality for problem $(\mathcal{P}^*_{\mathcal{M}})$. Set

$$\mathcal{F} : \mathcal{W} \to \overline{\mathbb{R}}, \qquad \mathcal{F}(q) = \frac{1}{2}\left\|A^*q + z\right\|^2_{\mathrm{L}^2} - \frac{1}{2}\left\|z\right\|^2_{\mathrm{L}^2}, \tag{2.5}$$

$$\mathcal{G} : \mathrm{C}_0(\Omega) \to \overline{\mathbb{R}}, \qquad \mathcal{G}(q) = I_{\{\|q\|_{\mathrm{C}_0} \le \alpha\}} := \begin{cases} 0 & \text{if } \|q\|_{\mathrm{C}_0} \le \alpha, \\ \infty & \text{if } \|q\|_{\mathrm{C}_0} > \alpha, \end{cases} \tag{2.6}$$

and $\Lambda : \mathcal{W} \to \mathrm{C}_0(\Omega)$ the injection given by the continuous embedding. The Fenchel conjugate of $\mathcal{F}$ is given by

$$\mathcal{F}^* : \mathcal{W}^* \to \overline{\mathbb{R}}, \qquad \mathcal{F}^*(u) = \frac{1}{2}\left\|A^{-1}u - z\right\|^2_{\mathrm{L}^2}. \tag{2.7}$$

The conjugate of $\mathcal{G}$ can be calculated as

$$\begin{aligned} \mathcal{G}^*(u) &= \sup_{q \in \mathrm{C}_0(\Omega)} \langle u, q\rangle_{\mathcal{M},\mathrm{C}_0} - I_{\{\|q\|_{\mathrm{C}_0} \le \alpha\}} = \sup_{\substack{q \in \mathrm{C}_0(\Omega),\\ \|q\|_{\mathrm{C}_0} \le \alpha}} \langle u, q\rangle_{\mathcal{M},\mathrm{C}_0} \\ &= \alpha \sup_{\substack{q \in \mathrm{C}_0(\Omega),\\ \|q\|_{\mathrm{C}_0} \le 1}} \langle u, q\rangle_{\mathcal{M},\mathrm{C}_0} = \alpha \left\|u\right\|_{\mathcal{M}}, \end{aligned} \tag{2.8}$$

and $\Lambda^* : \mathcal{M}(\Omega) \to \mathcal{W}^*$ is again the injection from the dual of $\mathrm{C}_0(\Omega)$ in $\mathcal{W}^*$.

It remains to verify the conditions of the Fenchel duality theorem. Since the norms in $\mathrm{L}^2(\Omega)$ and $\mathrm{C}_0(\Omega)$ are convex and lower semicontinuous, so are $\mathcal{F}$ and $\mathcal{G}$ (as indicator function of a convex set), which are also proper (*e.g.*, for $q = 0$, at which point $\mathcal{G}$ is continuous). In addition, Λ is a continuous linear operator, and so we have that

$$\min_{\substack{p \in \mathcal{W},\\ \|p\|_{\mathrm{C}_0} \le \alpha}} \frac{1}{2}\left\|A^*p + z\right\|^2_{\mathrm{L}^2} - \frac{1}{2}\left\|z\right\|^2_{\mathrm{L}^2} = \min_{u \in \mathcal{M}(\Omega)} \frac{1}{2}\left\|A^{-1}u - z\right\|^2_{\mathrm{L}^2} + \alpha\left\|-u\right\|_{\mathcal{M}}. \tag{2.9}$$

Introducing $y \in \mathrm{W}^{1,p}_0(\Omega)$ as the solution of $Ay = u \in \mathcal{M}(\Omega)$, we recover problem $(\mathcal{P}_{\mathcal{M}})$, and the relation (2.4) follows from the extremality relation (1.7). □

From this, we can derive the first order optimality conditions for problem $(\mathcal{P}^*_{\mathcal{M}})$:

Corollary 2.5. *Let $p^* \in \mathcal{W}$ be a solution of $(\mathcal{P}^*_{\mathcal{M}})$. Then there exists $\lambda^* \in \mathcal{M}(\Omega) \subset \mathcal{W}^*$ such that*

$$\begin{cases} \langle A^*p^* + z, A^*v\rangle_{\mathrm{L}^2} + \langle \lambda^*, v\rangle_{\mathcal{M},\mathrm{C}_0} = 0, \\ \langle \lambda^*, p - p^*\rangle_{\mathcal{M},\mathrm{C}_0} \le 0, \end{cases} \tag{2.10}$$

holds for all $v, p \in \mathcal{W}$ with $\|p\|_{\mathrm{C}_0} \le \alpha$. Moreover, the solution (p^*, λ^*) of (2.10) is unique.

Proof. By setting $\lambda^* = -u^*$ in the extremality relations (2.4), we immediately obtain (2.10) and existence of the Lagrange multiplier. Let (p_1, λ_1) and (p_2, λ_2) be two solutions of (2.10), and set $\delta p = p_1 - p_2$, $\delta\lambda = \lambda_1 - \lambda_2$. Setting $v = \delta p$, we have for all $v \in \mathcal{W}$

$$\langle A^*\delta p, A^*v\rangle_{\mathrm{L}^2} + \langle \delta\lambda, v\rangle_{\mathcal{M},\mathrm{C}_0} = 0, \tag{2.11}$$

$$\langle \delta\lambda, \delta p\rangle_{\mathcal{M},\mathrm{C}_0} \ge 0. \tag{2.12}$$

The choice $v = \delta p$ implies that

$$\left\|A^*\delta p\right\|^2_{\mathrm{L}^2} = -\langle \delta\lambda, \delta p\rangle_{\mathcal{M},\mathrm{C}_0} \le 0, \tag{2.13}$$

and hence $\delta p = 0$. By the assumptions on A^*, it follows that $\delta\lambda = 0$. □

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Splitting the first equation of (2.4) by introducing $y^* = A^*p + z$ yields

$$\begin{cases} Ay^* = u^*, \\ A^*p^* = y^* - z, \\ \quad 0 \geq \langle u^*, p^* - p\rangle_{\mathcal{M},\mathrm{C}_0} \ \text{ for all } p \in \mathcal{W},\ \|p\|_{\mathrm{C}_0} \leq \alpha. \end{cases} \tag{2.14}$$

Formally, this is the optimality system for $(\mathcal{P}_\mathcal{M})$, where $-p$ is the Lagrange multiplier for the constraint $Ay = u$, except that the non-reflexive Banach spaces $\mathcal{M}(\Omega)$ and $\mathcal{M}(\Omega)^*$ have been replaced by the Hilbert spaces $\mathcal{W}^*$ and $\mathcal{W}$, respectively.

From (2.14), we can deduce extra regularity for the Lagrange multiplier if the data is sufficiently smooth:

Corollary 2.6. *Let $1 \leq p < \frac{n}{n-1}$. If $z \in \mathrm{W}^{1,p}(\Omega)$, $\partial\Omega \in C^{2,1}$ and the coefficients of A are in $\mathrm{W}^{1,p}(\Omega)$ and $C^{2,1}$ for the principal part, then the optimal Lagrange multiplier for the state constraint in $(\mathcal{P}_\mathcal{M})$ satisfies $p^* \in \mathcal{W} \cap \mathrm{W}^{3,p}(\Omega)$.*

Proof. From Proposition 2.1 and $u^* \in \mathcal{M}(\Omega)$, we obtain that the optimal state satisfies $y^* \in \mathrm{W}_0^{1,p}(\Omega)$ for all $1 \leq p < \frac{n}{n-1}$. Hence, $A^*p^* = y^* - z$ yields $p^* \in \mathrm{W}^{3,p}(\Omega)$. □

We can also give the following structural information for the solution of problem $(\mathcal{P}_\mathcal{M})$:

Corollary 2.7. *Let u^* be the minimizer of $(\mathcal{P}_\mathcal{M})$ and p^* the minimizer of $(\mathcal{P}^*_\mathcal{M})$. Then, $u^* = u^*_+ - u^*_-$, where u^*_+ and u^*_- are positive measures with support:*

$$\operatorname{supp}(u^*_+) = \{x \in \Omega : p^*(x) = -\alpha\}, \tag{2.15}$$

$$\operatorname{supp}(u^*_-) = \{x \in \Omega : p^*(x) = \alpha\}. \tag{2.16}$$

This can be interpreted as a sparsity property: the optimal control u^* will be nonzero only on sets where the constraint on the dual variable p^* is active; hence the larger the penalty α, the smaller the support of the control.

Remark 2.8. If in addition, the minimizer satisfies $u^* \in \mathrm{L}^1(\Omega)$, it is the solution of the problem

$$\begin{cases} \min\limits_{u\in\mathrm{L}^1(\Omega)} \dfrac{1}{2}\|y - z\|^2_{\mathrm{L}^2} + \alpha\|u\|_{\mathrm{L}^1} \\ \text{s.t. } Ay = u. \end{cases} \tag{$\mathcal{P}_{\mathrm{L}^1}$}$$

This follows from the embedding of $\mathrm{L}^1(\Omega)$ into $\mathcal{M}(\Omega)$ and the fact that $\|v\|_\mathcal{M} = \|v\|_{\mathrm{L}^1}$ for $v \in \mathrm{L}^1(\Omega)$ (*cf.* [4], Chap. IV).

2.1.1. *Regularization of $(\mathcal{P}_\mathcal{M})$*

If we wish to avoid measures, we have to look for the minimizer in a stronger space than $\mathrm{L}^1(\Omega)$, *e.g.*, $\mathrm{L}^2(\Omega)$. Consider then the following regularized problem:

$$\begin{cases} \min\limits_{u\in\mathrm{L}^2(\Omega)} \dfrac{1}{2}\|y - z\|^2_{\mathrm{L}^2} + \alpha\|u\|_{\mathrm{L}^1} + \dfrac{\beta}{2}\|u\|^2_{\mathrm{L}^2} \\ \text{s.t. } Ay = u. \end{cases} \tag{$\mathcal{P}_{\mathrm{L}^1,\mathrm{L}^2}$}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Existence and uniqueness of a minimizer follows from standard arguments. Since $\mathrm{L}^2(\Omega)$ is reflexive, we can directly calculate the dual problem: Set

$$\mathcal{F} : \mathrm{L}^2(\Omega) \to \overline{\mathbb{R}}, \qquad \mathcal{F}(u) = \alpha \|u\|_{\mathrm{L}^1}, \tag{2.17}$$

$$\mathcal{G} : \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Omega) \to \overline{\mathbb{R}}, \qquad \mathcal{G}(u, y) = \frac{1}{2}\|y - z\|^2_{\mathrm{L}^2} + \frac{\beta}{2}\|u\|^2_{\mathrm{L}^2}, \tag{2.18}$$

$$\Lambda : \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Omega), \qquad \Lambda u = (u, A^{-1}u). \tag{2.19}$$

The Fenchel conjugate of $\mathcal{F}$ is again given by

$$\mathcal{F}^* : \mathrm{L}^2(\Omega) \to \overline{\mathbb{R}}, \qquad \mathcal{F}(\overline{u}) = I_{\{\|\overline{u}\|_{\mathrm{L}^\infty} \le \alpha\}}, \tag{2.20}$$

and we can calculate

$$\begin{cases} \mathcal{G}^* : \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Omega) \to \overline{\mathbb{R}}, \\ \mathcal{G}^*(\overline{u}, \overline{y}) = \frac{1}{2}\|\overline{y} + z\|^2_{\mathrm{L}^2} - \frac{1}{2}\|z\|^2_{\mathrm{L}^2} + \frac{1}{2\beta}\|\overline{u}\|^2_{\mathrm{L}^2}, \end{cases} \tag{2.21}$$

as well as

$$\Lambda^* : \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Omega) \to \mathrm{L}^2(\Omega), \quad \Lambda^*(\overline{u}, \overline{y}) = \overline{u} + A^{-*}\overline{y}. \tag{2.22}$$

Since $\mathcal{F}$ and $\mathcal{G}$ are convex and continuous operators (due to the continuous embedding of $\mathrm{L}^2(\Omega)$ into $\mathrm{L}^1(\Omega)$), and Λ is a continuous linear operator due to the well-posedness of the equality constraint, the Fenchel duality theorem yields the existence of a minimizer $(\overline{u}^*, \overline{y}^*) \in \mathrm{L}^2(\Omega) \times \mathrm{L}^2(\Omega)$ of the dual problem

$$\begin{cases} \min\limits_{(\overline{u},\overline{y}) \in \mathrm{L}^2 \times \mathrm{L}^2} \frac{1}{2}\|-\overline{y} + z\|^2_{\mathrm{L}^2} - \frac{1}{2}\|z\|^2_{\mathrm{L}^2} + \frac{1}{2\beta}\|-\overline{u}\|^2_{\mathrm{L}^2} \\ \text{s.t.} \quad \left\|\overline{u} + A^{-*}\overline{y}\right\|_{\mathrm{L}^\infty} \le \alpha, \end{cases} \tag{2.23}$$

where we can substitute $p = -A^{-*}\overline{y} \in \mathcal{W}$ and $q = \overline{u} - p \in \mathrm{L}^2(\Omega)$ to arrive at

$$\begin{cases} \min\limits_{(p,q) \in \mathcal{W} \times \mathrm{L}^2} \frac{1}{2}\|A^*p + z\|^2_{\mathrm{L}^2} - \frac{1}{2}\|z\|^2_{\mathrm{L}^2} + \frac{1}{2\beta}\|p + q\|^2_{\mathrm{L}^2} \\ \text{s.t.} \quad \|q\|_{\mathrm{L}^\infty} \le \alpha. \end{cases} \tag{$\mathcal{P}^*_{\mathrm{L}^1,\mathrm{L}^2}$}$$

In terms of p, the extremality relations linking the primal and dual minimizers u^* and (p^*, q^*) can be given as

$$\begin{cases} \langle u^*, v_1\rangle_{\mathrm{L}^2} = \langle A^*p^* + z, A^*v_1\rangle_{\mathrm{L}^2}, \\ \langle u^*, v_2\rangle_{\mathrm{L}^2} = \frac{1}{\beta}\langle q^* + p^*, v_2\rangle_{\mathrm{L}^2}, \\ 0 \ge \langle u^*, q^* - q\rangle_{\mathrm{L}^2}, \end{cases} \tag{2.24}$$

for all $v_1 \in \mathcal{W}$, $v_2, q \in \mathrm{L}^2(\Omega)$ with $\|q\|_{\mathrm{L}^\infty} \le \alpha$.

Note that now the Lagrange multiplier corresponding to the box constraint is in $\mathrm{L}^2(\Omega)$ (as opposed to $\mathcal{W}^*$). Problem ($\mathcal{P}^*_{\mathrm{L}^1,\mathrm{L}^2}$) therefore can be seen as a regularization of ($\mathcal{P}^*_{\mathcal{M}}$) by introducing a new variable $q = -p$ and treating this equality constraint by penalization (*cf.* also [11]). In Section 3 we will directly regularize the dual problem ($\mathcal{P}^*_{\mathcal{M}}$) by a Moreau-Yosida penalization with a parameter c, obtaining a regularized problem ($\mathcal{P}^*_{\mathcal{M},c}$). We will see that the Fenchel dual of ($\mathcal{P}^*_{\mathcal{M},c}$) is ($\mathcal{P}_{\mathrm{L}^1,\mathrm{L}^2}$) for an appropriate choice of the parameter c (see Rem. 3.2).

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

2.2. Controls in BV(Ω)

We now consider the optimal control problem

$$\begin{cases} \min\limits_{u\in \mathrm{BV}(\Omega)} \dfrac{1}{2}\|y-z\|^2_{\mathrm{L}^2} + \alpha\,\|u\|_{\mathrm{BV}} \\ \text{s.t. } Ay = u. \end{cases} \tag{$\mathcal{P}_{\mathrm{BV}}$}$$

Since the linear operator A^{-1} is injective on $\mathrm{BV}(\Omega) \subset \mathrm{L}^1(\Omega)$ by Proposition 2.1, the existence of a minimizer follows directly from [5], Theorem 2.1:

Proposition 2.9. *If $n = 2$, problem ($\mathcal{P}_{\mathrm{BV}}$) has a unique solution $(y^*, u^*) \in \mathrm{L}^2(\Omega) \times \mathrm{BV}(\Omega)$.*

Now let

$$\mathrm{H}^2_{\mathrm{div}}(\Omega) := \left\{ v \in \mathbb{L}^2(\Omega) : \operatorname{div} v \in \mathcal{W},\ v\cdot\nu = 0 \text{ on } \partial\Omega \right\}, \tag{2.25}$$

endowed with the norm $\|v\|^2_{\mathrm{H}^2_{\mathrm{div}}} := \|v\|^2_{\mathbb{L}^2} + \|\operatorname{div} v\|^2_{\mathcal{W}}$, and consider

$$\begin{cases} \min\limits_{p\in \mathrm{H}^2_{\mathrm{div}}(\Omega)} \dfrac{1}{2}\|A^*\operatorname{div} p + z\|^2_{\mathrm{L}^2} - \dfrac{1}{2}\|z\|^2_{\mathrm{L}^2} \\ \text{s.t. } \|p\|_{\mathbb{L}^\infty} \le \alpha. \end{cases} \tag{$\mathcal{P}^*_{\mathrm{BV}}$}$$

This problem has a solution, which, however, may not be unique. Set

$$\mathrm{H}^2_{\mathrm{div},0}(\Omega) = \left\{ v \in \mathrm{H}^2_{\mathrm{div}}(\Omega) : \operatorname{div} v = 0 \right\} \tag{2.26}$$

and let $\mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ the orthogonal complement in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$. Then we can show the following:

Theorem 2.10. *Problem ($\mathcal{P}^*_{\mathrm{BV}}$) has a solution $p^* \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$. Moreover, there exists a unique $q^* \in \mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ such that all such solutions satisfy $p^* \in \{q^*\} + \mathrm{H}^2_{\mathrm{div},0}(\Omega)$.*

Proof. Consider again a minimizing sequence $\{p_n\}_{n\in\mathbb{N}} \subset \mathrm{H}^2_{\mathrm{div}}(\Omega)$. The $\mathbb{L}^2(\Omega)$-norm of p_n is bounded *via* the box constraints, and the data fit term gives a bound on the $\mathcal{W}$-norm of $(\operatorname{div} p_n)$; together, this yields that the $\mathrm{H}^2_{\mathrm{div}}(\Omega)$-norm of p_n is bounded. We can therefore extract a subsequence weakly converging in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$, and existence of the minimizer follows from the same arguments as in the proof of Theorem 2.3.

Since $\mathrm{H}^2_{\mathrm{div},0}(\Omega)$ is a closed subspace of $\mathrm{H}^2_{\mathrm{div}}(\Omega)$, it holds that

$$\mathrm{H}^2_{\mathrm{div}}(\Omega) = \mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp \oplus \mathrm{H}^2_{\mathrm{div},0}(\Omega), \tag{2.27}$$

and that div is injective on $\mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ by construction. Therefore, $\frac{1}{2}\|A^*\operatorname{div} p + z\|^2_{\mathrm{L}^2}$ is strictly convex on $\mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$, so that ($\mathcal{P}^*_{\mathrm{BV}}$) has a unique minimizer $q^* \in \mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ there. On the other hand, given $p \in \mathrm{H}^2_{\mathrm{div},0}(\Omega)$ with $\|q^* + p\|_{\mathbb{L}^\infty} \le \alpha$, we find that $q^* + p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ is also a minimizer. □

Theorem 2.11. *The dual of ($\mathcal{P}^*_{\mathrm{BV}}$) is ($\mathcal{P}_{\mathrm{BV}}$), and the solutions $u^* \in \mathrm{BV}(\Omega)$ of ($\mathcal{P}_{\mathrm{BV}}$) and $p^* \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ of ($\mathcal{P}^*_{\mathrm{BV}}$) are related by*

$$\begin{cases} \langle -u^*, v\rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*,\mathrm{H}^2_{\mathrm{div}}} = \langle A^*(-\operatorname{div})p^* + z, A^*v\rangle_{\mathrm{L}^{2*},\mathrm{L}^2}, \\ 0 \ge \langle (-\operatorname{div})^* u^*, p - p^*\rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*,\mathrm{H}^2_{\mathrm{div}}}, \end{cases} \tag{2.28}$$

for all $v \in \mathcal{W}$, $p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ with $\|p\|_{\mathbb{L}^\infty} \le \alpha$.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Proof. We apply Fenchel duality. Setting

$$\mathcal{F} : \mathrm{H}^2_{\mathrm{div}}(\Omega) \to \overline{\mathbb{R}}, \quad \mathcal{F}(q) = I_{\{\|q\|_{\mathbb{L}^\infty} \leq \alpha\}}, \tag{2.29}$$

$$\mathcal{G} : \mathcal{W} \to \overline{\mathbb{R}}, \qquad \mathcal{G}(q) = \frac{1}{2} \|A^* q - z\|^2_{\mathrm{L}^2} - \frac{1}{2} \|z\|^2_{\mathrm{L}^2}, \tag{2.30}$$

$$\Lambda : \mathrm{H}^2_{\mathrm{div}}(\Omega) \to \mathcal{W}, \qquad \Lambda q = -\operatorname{div} q, \tag{2.31}$$

problem $(\mathcal{P}^*_{\mathrm{BV}})$ can be written as

$$\min_{p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)} \mathcal{F}(p) + \mathcal{G}(\Lambda p). \tag{2.32}$$

The Fenchel conjugate of $\mathcal{G}$ is given by

$$\mathcal{G}^* : \mathrm{H}^2_{\mathrm{div}}(\Omega)^* \to \overline{\mathbb{R}}, \qquad \mathcal{G}^*(v) = \frac{1}{2} \left\|A^{-1} v + z\right\|^2_{\mathrm{L}^2}, \tag{2.33}$$

and the adjoint of Λ is

$$\Lambda^* : \mathcal{W}^* \to \mathrm{H}^2_{\mathrm{div}}(\Omega)^*, \quad \Lambda^* v = (-\operatorname{div})^* v. \tag{2.34}$$

It remains to calculate $\mathcal{F}^* : \mathrm{H}^2_{\mathrm{div}}(\Omega)^* \to \overline{\mathbb{R}}$. We have, as in (2.8), that

$$\mathcal{F}^*(v) = \alpha \sup_{\substack{q \in \mathrm{H}^2_{\mathrm{div}}(\Omega), \\ \|q\|_{\mathbb{L}^\infty} \leq 1}} \langle v, q \rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*, \mathrm{H}^2_{\mathrm{div}}}. \tag{2.35}$$

By standard arguments, we can show that $(\mathrm{C}^\infty_0(\Omega))^n$ is dense in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$ (*cf.* [13], p. 26, and [2]), so that we can equivalently write

$$\mathcal{F}^*(v) = \alpha \sup_{\substack{q \in (\mathrm{C}^\infty_0(\Omega))^n, \\ \|q\|_{(\mathrm{C}_0)^n} \leq 1}} \langle v, q \rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*, \mathrm{H}^2_{\mathrm{div}}}, \tag{2.36}$$

and thus

$$\mathcal{F}^*((-\operatorname{div})^* u) = \alpha \sup_{\substack{q \in (\mathrm{C}^\infty_0(\Omega))^n, \\ \|q\|_{(\mathrm{C}_0)^n} \leq 1}} \langle u, -\operatorname{div} q \rangle_{\mathcal{W}^*, \mathcal{W}} = \alpha \|u\|_{\mathrm{BV}}, \tag{2.37}$$

which is finite if and only if $u \in \mathrm{BV}(\Omega)$.

Again, $\mathcal{F}$ is convex and lower semicontinuous, $\mathcal{G}$ is convex and continuous on $\mathcal{W}$, and Λ is a continuous linear operator. The Fenchel duality theorem thus yields the duality of $(\mathcal{P}^*_{\mathrm{BV}})$ and

$$\max_{u \in \mathcal{W}^*} -\mathcal{F}^*(\Lambda^* u) - \mathcal{G}^*(-u) = -\inf_{u \in \mathrm{BV}(\Omega)} \alpha \|u\|_{\mathrm{BV}} + \frac{1}{2} \left\|A^{-1} u - z\right\|^2_{\mathrm{L}^2}, \tag{2.38}$$

by which we recover $(\mathcal{P}_{\mathrm{BV}})$. The relations (2.28) are once more the explicit form of the extremality relations (1.7). □

From this, we also obtain the first order necessary optimality conditions:

Corollary 2.12. *Let $p^* \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ be a solution of $(\mathcal{P}^*_{\mathrm{BV}})$. Then there exists $\lambda^* \in \mathrm{H}^2_{\mathrm{div}}(\Omega)^*$ such that*

$$\begin{cases} \langle A^*(-\operatorname{div}) p^* + z, A^*(-\operatorname{div}) w \rangle_{\mathrm{L}^2} + \langle \lambda^*, v \rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*, \mathrm{H}^2_{\mathrm{div}}} = 0, \\ \langle \lambda^*, p - p^* \rangle_{{\mathrm{H}^2_{\mathrm{div}}}^*, \mathrm{H}^2_{\mathrm{div}}} \leq 0, \end{cases} \tag{2.39}$$

holds for all $w, p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ with $\|p\|_{\mathbb{L}^\infty} \leq \alpha$. Moreover, the solution (p^*, λ^*) of (2.10) is unique in $\mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp \times \mathrm{H}^2_{\mathrm{div}}(\Omega)^*$.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Proof. For $w \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$, we insert $v = -\operatorname{div} w \in \mathcal{W}$ in the first relation of (2.28), and set $\lambda^* := (-\operatorname{div})^* u^* \in \mathrm{H}^2_{\mathrm{div}}(\Omega)^*$. Proceeding as in the proof of Corollary 2.5, we deduce that λ^* solving (2.39) is unique. Since the operator $\langle A^*(-\operatorname{div})\cdot, A^*(-\operatorname{div})\cdot\rangle_{\mathrm{L}^2}$ is an inner product on $\mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ by construction, we obtain a unique $q^* \in \mathrm{H}^2_{\mathrm{div},0}(\Omega)^\perp$ such that all solutions p^* of (2.28) satisfy $p^* \in \{q^*\} + \mathrm{H}^2_{\mathrm{div},0}(\Omega)$. □

Note that this implies that any solution p^* of problem ($\mathcal{P}^*_{\mathrm{BV}}$) will yield the same (unique) solution u^* of ($\mathcal{P}_{\mathrm{BV}}$) when calculated *via* the extremality relation (2.28).

Remark 2.13. In order to obtain a unique solution for problem ($\mathcal{P}^*_{\mathrm{BV}}$), we can consider the following regularized problem for $\beta > 0$:

$$\begin{cases} \min\limits_{p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)} \dfrac{1}{2} \left\| A^* \operatorname{div} p + z \right\|^2_{\mathrm{L}^2} - \dfrac{1}{2} \|z\|^2_{\mathrm{L}^2} + \dfrac{\beta}{2} \|p\|^2_{\mathbb{L}^2} \\ \text{s.t.} \quad \|p\|_{\mathbb{L}^\infty} \leq \alpha. \end{cases} \tag{$\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2}$}$$

This can be expressed as a regularization of the primal problem ($\mathcal{P}_{\mathrm{BV}}$) as well: Setting

$$\mathcal{F}(q) = I_{\{\|q\|_{\mathbb{L}^\infty} \leq \alpha\}} + \frac{\beta}{2} \|q\|^2_{\mathbb{L}^2} \tag{2.40}$$

and Λ as in Theorem 2.11, we find that

$$\mathcal{F}^*(\Lambda^* u) = \int_\Omega \varphi(\nabla u)\, \mathrm{d}x, \tag{2.41}$$

where

$$\varphi(\vec{v})(x) = \begin{cases} \frac{1}{2\beta} |\vec{v}(x)|^2 & \text{if } |\vec{v}(x)| < \alpha\beta, \\ \alpha |\vec{v}(x)| - \frac{\alpha\beta}{2} & \text{if } |\vec{v}(x)| \geq \alpha\beta. \end{cases} \tag{2.42}$$

In the primal problem, the additional predual $\mathbb{L}^2(\Omega)$-regularization essentially results in locally replacing the $\mathrm{BV}(\Omega)$-term with a $\mathrm{H}^1(\Omega)$-penalty in a small neighborhood of the origin (*cf.* [7]).

We can finally give some structural information on the optimal control in $\mathrm{BV}(\Omega)$:

Corollary 2.14. *Let u^* be a minimizer of ($\mathcal{P}_{\mathrm{BV}}$). Then the following holds for any $p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$, $p \geq 0$:*

$$\langle (-\operatorname{div})^* u^*, p \rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*, \mathrm{H}^2_{\mathrm{div}}} = 0 \quad \textit{if } \operatorname{supp} p \subset \bigcap_{i=1}^n \{x : |p_i^*(x)| < \alpha\}, \tag{2.43}$$

$$\langle (-\operatorname{div})^* u^*, p \rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*, \mathrm{H}^2_{\mathrm{div}}} \geq 0 \quad \textit{if } \operatorname{supp} p \subset \bigcup_{i=1}^n \{x : p_i^*(x) = \alpha\}, \tag{2.44}$$

$$\langle (-\operatorname{div})^* u^*, p \rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*, \mathrm{H}^2_{\mathrm{div}}} \leq 0 \quad \textit{if } \operatorname{supp} p \subset \bigcup_{i=1}^n \{x : p_i^*(x) = -\alpha\}. \tag{2.45}$$

Again, this can be interpreted as a sparsity condition on the gradient of the control: the optimal control u^* will be piecewise constant on sets where the constraints on the dual variable p^* are inactive.

3. Solution of the optimality systems

This section, we present a method for the numerical solution of the optimality systems (2.10) and (2.39). For this, we need to deal with the fact that the Lagrange multiplier corresponding to the box constraints is in general only in $\mathcal{W}^*$ (and $\mathrm{H}^2_{\mathrm{div}}(\Omega)^*$, respectively). We introduce therefore regularized versions of (2.10) and (2.39) which can be solved using a semismooth Newton method having superlinear convergence.

Again, we first treat the solution of (2.10), and discuss the corresponding results and algorithm for problem (2.39) more briefly in Section 3.3.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

3.1. **Regularization of box constraints**

In order to obtain Lagrange multipliers in $\mathrm{L}^2(\Omega)$, we introduce the Moreau-Yosida regularization of problem ($\mathcal{P}^*_{\mathcal{M}}$) for $c > 0$:

$$\min_{p\in\mathcal{W}} \frac{1}{2}\left\|A^*p+z\right\|^2_{\mathrm{L}^2} - \frac{1}{2}\left\|z\right\|^2_{\mathrm{L}^2} + \frac{1}{2c}\left\|\max(0,c(p-\alpha))\right\|^2_{\mathrm{L}^2} + \frac{1}{2c}\left\|\min(0,c(p+\alpha))\right\|^2_{\mathrm{L}^2}, \tag{$\mathcal{P}^*_{\mathcal{M},c}$}$$

where the max and min are taken pointwise in Ω. This is equivalent to the regularization in problem ($\mathcal{P}^*_{\mathrm{L}^1,\mathrm{L}^2}$) (*cf.* Rem. 3.2) but more amenable to solution by a semismooth Newton method. Existence and uniqueness of a minimizer is directly deduced from lower semicontinuity and strict convexity of the functional. The corresponding optimality system is given by

$$\begin{cases} \langle A^*p_c, A^*v\rangle_{\mathrm{L}^2} + \langle z, A^*v\rangle_{\mathrm{L}^2} + \langle \lambda_c, v\rangle_{\mathcal{W}^*,\mathcal{W}} = 0, \\ \lambda_c = \max(0, c(p_c-\alpha)) + \min(0, c(p_c+\alpha)), \end{cases} \tag{3.1}$$

for all $v \in \mathcal{W}$, where the Lagrange multiplier satisfies $\lambda_c \in \mathrm{W}^{1,\infty}(\Omega)$.

First, we address the convergence of the solutions of (3.1) as c tends to infinity:

Theorem 3.1. *Let $(p_c, \lambda_c) \in \mathcal{W}\times\mathcal{W}^*$ be the solution of (3.1) for given $c > 0$, and $(p^*, \lambda^*) \in \mathcal{W}\times\mathcal{W}^*$ be the unique solution of (2.10). Then we have as $c \to \infty$:*

$$p_c \rightharpoonup p^* \quad \text{in } \mathcal{W}, \tag{3.2}$$

$$\lambda_c \rightharpoonup \lambda^* \quad \text{in } \mathcal{W}^*. \tag{3.3}$$

Proof. From the optimality conditions (3.1), we have that pointwise in $x \in \Omega$

$$\lambda_c p_c = \max(0, c(p_c-\alpha))p_c + \min(0, c(p_c+\alpha))p_c = \begin{cases} c(p_c-\alpha)p_c, & p_c \geq \alpha, \\ 0, & |p_c| < \alpha, \\ c(p_c+\alpha)p_c, & p_c \leq -\alpha, \end{cases} \tag{3.4}$$

and hence that

$$\langle \lambda_c, p_c\rangle_{\mathrm{L}^2} \geq \frac{1}{c}\left\|\lambda_c\right\|^2_{\mathrm{L}^2}. \tag{3.5}$$

Inserting p_c in (3.1) yields

$$\left\|A^*p_c\right\|^2_{\mathrm{L}^2} + \frac{1}{c}\left\|\lambda_c\right\|^2_{\mathrm{L}^2} \leq \left\|A^*p_c\right\|_{\mathrm{L}^2}\left\|z\right\|_{\mathrm{L}^2}, \tag{3.6}$$

and we deduce that $\left\|A^*p_c\right\|_{\mathrm{L}^2} \leq \left\|z\right\|_{\mathrm{L}^2}$, as well as

$$\begin{aligned} \left\|\lambda_c\right\|_{\mathcal{W}^*} &= \sup_{\substack{v\in\mathcal{W},\\ \|v\|_{\mathcal{W}}\leq 1}} \langle \lambda_c, v\rangle_{\mathcal{W}^*,\mathcal{W}} \leq \sup_{\substack{v\in\mathcal{W},\\ \|v\|_{\mathcal{W}}\leq 1}} \left[\langle A^*p_c, A^*v\rangle_{\mathrm{L}^2} + \langle z, A^*v\rangle_{\mathrm{L}^2}\right] \\ &\leq 2 \sup_{\substack{v\in\mathcal{W},\\ \|v\|_{\mathcal{W}}\leq 1}} \left\|A^*v\right\|_{\mathrm{L}^2}\left\|z\right\|_{\mathrm{L}^2} =: K < \infty. \end{aligned} \tag{3.7}$$

Thus, (p_c, λ_c) is uniformly bounded in $\mathcal{W}\times\mathcal{W}^*$, so that we can deduce the existence of a $(\tilde{p}, \tilde{\lambda}) \in \mathcal{W}\times\mathcal{W}^*$ such that

$$(p_c, \lambda_c) \rightharpoonup (\tilde{p}, \tilde{\lambda}) \quad \text{in } \mathcal{W}\times\mathcal{W}^*. \tag{3.8}$$

Passing to the limit in (3.1), we obtain

$$\langle A^*\tilde{p}, A^*v\rangle_{\mathrm{L}^2} + \langle z, A^*v\rangle_{\mathrm{L}^2} + \left\langle \tilde{\lambda}, v\right\rangle_{\mathcal{W}^*,\mathcal{W}} = 0 \qquad \text{for all } v \in \mathcal{W}. \tag{3.9}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

We next verify the feasibility of $\tilde{p}$. By pointwise inspection similar to (3.5), we obtain that

$$\frac{1}{c}\,\|\lambda_c\|_{\mathrm{L}^2}^2 = c\,\|\max(0, p_c - \alpha)\|_{\mathrm{L}^2}^2 + c\,\|\min(0, p_c + \alpha)\|_{\mathrm{L}^2}^2\,. \tag{3.10}$$

From (3.6), we have that $\frac{1}{c}\,\|\lambda_c\|_{\mathrm{L}^2}^2 \leq \|z\|_{\mathrm{L}^2}^2$, so that

$$\|\max(0, p_c - \alpha)\|_{\mathrm{L}^2}^2 \leq \frac{1}{c}\,\|z\|_{\mathrm{L}^2}^2 \to 0, \tag{3.11}$$

$$\|\min(0, p_c + \alpha)\|_{\mathrm{L}^2}^2 \leq \frac{1}{c}\,\|z\|_{\mathrm{L}^2}^2 \to 0 \tag{3.12}$$

holds for $c \to \infty$. Since $p_c \to \tilde{p}$ strongly in $\mathrm{L}^2(\Omega)$, this implies that

$$-\alpha \leq \tilde{p}(x) \leq \alpha \quad \text{ for all } x \in \Omega. \tag{3.13}$$

It remains to pass to the limit in the second equation of (2.10). First, optimality of p_c yields that

$$\frac{1}{2}\,\|A^* p_c + z\|_{\mathrm{L}^2}^2 \leq \frac{1}{2}\,\|A^* p + z\|_{\mathrm{L}^2}^2 \tag{3.14}$$

holds for all feasible $p \in \mathcal{W}$. Therefore, we have that

$$\limsup_{c\to\infty} \frac{1}{2}\,\|A^* p_c + z\|_{\mathrm{L}^2}^2 \leq \frac{1}{2}\,\|A^* \tilde{p} + z\|_{\mathrm{L}^2}^2 \tag{3.15}$$

and thus $p_c \to \tilde{p}$ strongly in $\mathcal{W}$. Now observe that

$$\langle \lambda_c, p - p_c\rangle_{\mathcal{W}^*,\mathcal{W}} = \langle \max(0, c(p_c - \alpha)), p - p_c\rangle_{\mathcal{W}^*,\mathcal{W}} + \langle \min(0, c(p_c + \alpha)), p - p_c\rangle_{\mathcal{W}^*,\mathcal{W}} \leq 0 \tag{3.16}$$

holds for all $p \in \mathcal{W}$ with $\|p\|_{\mathrm{C}_0} \leq \alpha$, and thus

$$\left\langle \tilde{\lambda}, p - \tilde{p} \right\rangle_{\mathcal{W}^*,\mathcal{W}} \leq 0 \tag{3.17}$$

is satisfied for all $p \in \mathcal{W}$ with $\|p\|_{\mathrm{C}_0} \leq \alpha$. Therefore, $(\tilde{p}, \tilde{\lambda}) \in \mathcal{W} \times \mathcal{W}^*$ satisfies (2.10), and since the solution of (2.10) is unique, $\tilde{p} = p^*$ and $\tilde{\lambda} = \lambda^*$ follows. □

Remark 3.2. We can also relate the Moreau-Yosida regularization to the regularized problem ($\mathcal{P}^*_{\mathrm{L}^1,\mathrm{L}^2}$) *via* the primal problem: we proceed by calculating the dual of ($\mathcal{P}^*_{\mathcal{M},c}$) as in the proof of Theorem 2.4, defining $\mathcal{G} : \mathrm{C}_0(\Omega) \to \overline{\mathbb{R}}$,

$$\mathcal{G}(q) = \frac{1}{2c}\,\|\max(0, c(q - \alpha))\|_{\mathrm{L}^2}^2 + \frac{1}{2c}\,\|\min(0, c(q + \alpha))\|_{\mathrm{L}^2}^2\,. \tag{3.18}$$

To calculate the Fenchel conjugate, we use (1.5), which in this case implies that

$$u = \max(0, c(q - \alpha)) + \min(0, c(q + \alpha)) \tag{3.19}$$

has to hold for the primal variable $u \in M$. If $u(x) > 0$, the right hand side has to be positive as well, which implies $u(x) = c(q(x) - \alpha)$ and hence $q(x) = \frac{1}{c}u(x) + \alpha$. Similarly, $u(x) < 0$ yields $q(x) = \frac{u}{c}(x) - \alpha$.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

For $u(x) = 0$, $-\alpha < p(x) < \alpha$ holds. Substituting in the definition of $\mathcal{G}^*$, we have that

$$\begin{aligned}\mathcal{G}^*(u) &= \int_{\{u>0\}} u(x)\left(\frac{1}{c}u(x)+\alpha\right) - \frac{1}{2c}\max(0,u(x))^2\,\mathrm{d}x + \int_{\{u<0\}} u(x)\left(\frac{1}{c}u(x)-\alpha\right) - \frac{1}{2c}\min(0,u(x))^2\,\mathrm{d}x \\ &= \frac{1}{2c}\|u\|^2_{\mathrm{L}^2} + \alpha\|u\|_{\mathrm{L}^1}\,, \end{aligned} \tag{3.20}$$

which is finite if and only if $u \in \mathrm{L}^2(\Omega)$. Setting $\beta := \frac{1}{c}$, we arrive at problem $(\mathcal{P}_{\mathrm{L}^1,\mathrm{L}^2})$. Since both regularized problems are posed in reflexive Hilbert spaces, they are equivalent.

3.2. **Semismooth Newton method**

The regularized optimality system (3.1) can be solved efficiently using a semismooth Newton method (*cf.* [8,14]), which is superlinearly convergent. For this purpose, we consider (3.1) as a nonlinear equation $F(p) = 0$ for $F : \mathcal{W} \to \mathcal{W}^*$, defined for $v \in \mathcal{W}$ by

$$\langle F(p), v\rangle_{\mathcal{W}^*,\mathcal{W}} := \langle A^*p + z, v\rangle_{\mathrm{L}^2} + \langle \max(0, c(p-\alpha)) + \min(0, c(p+\alpha)), v\rangle_{\mathrm{L}^2}\,. \tag{3.21}$$

It is known (*cf.*, *e.g.*, [9], Ex. 8.14) that the projection operator

$$P_\alpha(p) := \max(0,(p-\alpha)) + \min(0,(p+\alpha)) \tag{3.22}$$

is semismooth from $\mathrm{L}^q(\Omega)$ to $\mathrm{L}^p(\Omega)$, if and only if $q > p$, and has as Newton-derivative

$$\partial_N P_\alpha(p)h := h\chi_{\{|p|>\alpha\}} = \begin{cases} h(x) & \text{if } |p(x)| > \alpha, \\ 0 & \text{if } |p(x)| \le \alpha. \end{cases} \tag{3.23}$$

Since Fréchet-differentiable functions and sums of semismooth functions are semismooth (with canonical Newton-derivatives), we find that F is semismooth, and that its Newton-derivative is defined by

$$\langle \partial_N F(p)h, v\rangle_{\mathcal{W}^*,\mathcal{W}} = \langle A^*h, A^*v\rangle_{\mathrm{L}^2} + c\left\langle h\chi_{\{|p|>\alpha\}}, v\right\rangle_{\mathrm{L}^2}\,, \tag{3.24}$$

for all $v \in \mathcal{W}$.

A semismooth Newton step consists in solving for p^{k+1} the equation

$$\partial_N F(p^k)(p^{k+1} - p^k) = -F(p^k). \tag{3.25}$$

Defining the active and inactive sets

$$\mathcal{A}_k^+ := \left\{x \in \Omega : p^k(x) > \alpha\right\}, \quad \mathcal{A}_k^- := \left\{x \in \Omega : p^k(x) < -\alpha\right\}, \quad \mathcal{A}_k := \mathcal{A}_k^+ \cup \mathcal{A}_k^-, \tag{3.26}$$

step (3.25) can be written explicitly as finding $p^{k+1} \in \mathcal{W}$ such that

$$\left\langle A^*p^{k+1}, A^*v\right\rangle_{\mathrm{L}^2} + c\left\langle p^{k+1}\chi_{\mathcal{A}_k}, v\right\rangle_{\mathrm{L}^2} = -\left\langle z, A^*v\right\rangle_{\mathrm{L}^2} + c\alpha\left\langle \chi_{\mathcal{A}_k^+} - \chi_{\mathcal{A}_k^-}, v\right\rangle_{\mathrm{L}^2} \tag{3.27}$$

for all $v \in \mathcal{W}$. The resulting semismooth Newton method is given as Algorithm 1.

Theorem 3.3. *If $\|p_c - p^0\|_{\mathcal{W}}$ is sufficiently small, the iterates p^k of Algorithm 1 converge superlinearly in $\mathcal{W}$ to the solution p_c of (3.1) as $k \to \infty$.*

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Algorithm 1 Semismooth Newton method for (3.1)

1: Set $k = 0$, Choose $p^0 \in \mathcal{W}$
2: **repeat**
3: Set $\mathcal{A}_k^+ = \{x \in \Omega : p^k(x) > \alpha\}$, $\mathcal{A}_k^- = \{x \in \Omega : p^k(x) < -\alpha\}$, $\mathcal{A}_k = \mathcal{A}_k^+ \cup \mathcal{A}_k^-$
4: Solve for $p^{k+1} \in \mathcal{W}$:

$$\left\langle A^* p^{k+1}, A^* v \right\rangle_{\mathrm{L}^2} + c \left\langle p^{k+1} \chi_{\mathcal{A}_k}, v \right\rangle_{\mathrm{L}^2} = - \left\langle z, A^* v \right\rangle_{\mathrm{L}^2} + c\alpha \left\langle \chi_{\mathcal{A}_k^+} - \chi_{\mathcal{A}_k^-}, v \right\rangle_{\mathrm{L}^2}$$

for all $v \in \mathcal{W}$
5: Set $k = k + 1$
6: **until** $(\mathcal{A}_k^+ = \mathcal{A}_{k-1}^+)$ and $(\mathcal{A}_k^- = \mathcal{A}_{k-1}^-)$

Proof. Since F is semismooth, it suffices to show that $(\partial_N F)^{-1}$ is uniformly bounded. Let $g \in \mathcal{W}^*$ be given. Due to the assumptions on A^*, the Riesz representation theorem ensures the existence of a unique $\varphi \in \mathcal{W}$ such that

$$\left\langle A^* \varphi, A^* v \right\rangle_{\mathrm{L}^2} + c \left\langle \chi_{\mathcal{A}} \varphi, v \right\rangle_{\mathrm{L}^2} = \left\langle g, v \right\rangle_{\mathcal{W}^*, \mathcal{W}} \tag{3.28}$$

holds for all $v \in \mathcal{W}$, independent of $\mathcal{A}$. Furthermore, φ satisfies

$$\|\varphi\|_{\mathcal{W}}^2 \leq C \|g\|_{\mathcal{W}^*}^2 , \tag{3.29}$$

with a constant C depending only on A and Ω, giving the desired uniform bound. The superlinear convergence now follows from standard results (*e.g.*, [9], Thm. 8.16). □

The termination criterion in Algorithm 1, step 6, can be justified as follows:

Proposition 3.4. *If $\mathcal{A}_{k+1}^+ = \mathcal{A}_k^+$ and $\mathcal{A}_{k+1}^- = \mathcal{A}_k^-$ holds, then p^{k+1} satisfies $F(p^{k+1}) = 0$.*

Proof. Since the solution of (3.27) is unique for fixed $\mathcal{A}_k^+, \mathcal{A}_k^-$, we have that $p^{k+1} = p^k$. Thus, setting $\mathcal{A}_{k+1}^+ = \mathcal{A}_k^+$ and $\mathcal{A}_{k+1}^- = \mathcal{A}_k^-$ in (3.27) and noting that

$$c\chi_{\mathcal{A}_{k+1}^+} p^{k+1} - c\alpha\chi_{\mathcal{A}_{k+1}^+} = \max(0, c(p^{k+1} - \alpha)), \tag{3.30}$$

we see that (3.27) is equivalent to $F(p^{k+1}) = 0$. It follows that p^{k+1} is a solution of (3.1). □

3.3. Controls in BV(Ω)

The arguments above rely on the fact that the term $\|A^* p\|_{\mathrm{L}^2}$ in the functional is an equivalent norm on $\mathcal{W}$. For the problem in $\mathrm{BV}(\Omega)$, the corresponding term $\|A^* \operatorname{div} p\|_{\mathrm{L}^2}$ is only a seminorm on $\mathrm{H}^2_{\mathrm{div}}(\Omega)$, and we need to add additional regularization. Since furthermore $\mathrm{H}^2_{\mathrm{div}}(\Omega)$ does not embed into $\mathbb{L}^q$ for $q > 2$, we set $\mathcal{H} := \mathrm{H}^2_{\mathrm{div}}(\Omega) \cap \mathcal{W}^n$ and consider

$$\min_{p \in \mathcal{H}} \frac{1}{2} \|A^* \operatorname{div} p + z\|_{\mathrm{L}^2}^2 + \frac{\beta}{2} \|p\|_{\mathcal{W}^n}^2 - \frac{1}{2} \|z\|_{\mathrm{L}^2}^2 + \frac{1}{2c} \|\max(0, c(p - \alpha))\|_{\mathbb{L}^2}^2 + \frac{1}{2c} \|\min(0, c(p + \alpha))\|_{\mathbb{L}^2}^2 \quad (\mathcal{P}^*_{\mathrm{BV},c})$$

with the corresponding optimality system

$$\begin{cases} \left\langle A^*(-\operatorname{div}) p_c + z, A^*(-\operatorname{div}) v \right\rangle_{\mathrm{L}^2} + \beta \left\langle \Delta p_c, \Delta v \right\rangle_{\mathbb{L}^2} + \left\langle \lambda_c, v \right\rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*, \mathrm{H}^2_{\mathrm{div}}} = 0, \\ \lambda_c = \max(0, c(p_c - \alpha)) + \min(0, c(p_c + \alpha)), \end{cases} \tag{3.31}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Algorithm 2 Semismooth Newton method for (3.31)

1: Set $k = 0$, Choose $p^0 \in \mathcal{H}$
2: **repeat**
3: Set for $i = 1, \dots, n$

$$\mathcal{A}^+_{i,k} = \left\{x \in \Omega : p_i^k(x) > \alpha\right\}, \quad \mathcal{A}^-_{i,k} = \left\{x \in \Omega : p_i^k(x) < -\alpha\right\}, \quad \mathcal{A}_{i,k} = \mathcal{A}^+_{i,k} \cup \mathcal{A}^-_{i,k}$$

4: Solve for $p^{k+1} \in \mathcal{H}$:

$$\left\langle A^* \operatorname{div} p^{k+1}, A^* \operatorname{div} v\right\rangle_{\mathrm{L}^2} + \beta \sum_{i=1}^n \left\langle \Delta p_i^{k+1}, \Delta v_i\right\rangle_{\mathrm{L}^2} + c \sum_{i=1}^n \left\langle p_i^{k+1} \chi_{\mathcal{A}_{k,i}}, v_i\right\rangle_{\mathrm{L}^2}$$
$$= \left\langle z, A^* \operatorname{div} v\right\rangle_{\mathrm{L}^2} + c\alpha \sum_{i=1}^n \left\langle \chi_{\mathcal{A}^+_{i,k}} - \chi_{\mathcal{A}^-_{i,k}}, v_i\right\rangle_{\mathrm{L}^2}$$

for all $v \in \mathcal{H}$
5: Set $k = k + 1$
6: **until** $(\mathcal{A}^+_{i,k} = \mathcal{A}^+_{i,k-1})$ and $(\mathcal{A}^-_{i,k} = \mathcal{A}^-_{i,k-1})$ for all $i = 1, \dots, n$

for all $v \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$, where Δ denotes the componentwise Laplacian with homogeneous Dirichlet boundary conditions, and the max, min are understood to act componentwise. (Here and below, α stands for the vector $(\alpha, \dots, \alpha) \in \mathbb{R}^n$.)

The convergence for $\beta \to 0$ is impeded by the fact that no unique candidate for the limit exists. We can, however, show convergence for the corresponding regularization of problem ($\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2}$), and the proof is given in the appendix. In the following, we will consider only the solution of problem ($\mathcal{P}^*_{\mathrm{BV},c}$).

We once more formulate the optimality system (3.31) as a semismooth operator equation $G(p) = 0$ for $G : \mathcal{H} \to \mathcal{H}^*$,

$$\langle G(p), v\rangle_{\mathcal{H}^*,\mathcal{H}} := \langle A^*(-\operatorname{div})p_c + z, A^*(-\operatorname{div})v\rangle_{\mathrm{L}^2} + \beta \langle \Delta p_c, \Delta v\rangle_{\mathbb{L}^2} + \langle \max(0, c(p-\alpha)) + \min(0, c(p+\alpha)), v\rangle_{\mathrm{L}^2} \tag{3.32}$$

for all $v \in \mathcal{H}$ with Newton derivative

$$\langle \partial_N G(p)h, v\rangle_{\mathcal{H}^*,\mathcal{H}} = \langle A^*(-\operatorname{div})h, A^*(-\operatorname{div})v\rangle_{\mathrm{L}^2} + \beta \langle \Delta h, \Delta v\rangle_{\mathrm{L}^2} + c \left\langle h\chi_{\{|p|>\alpha\}}, v\right\rangle_{\mathrm{L}^2} \tag{3.33}$$

for all $v \in \mathcal{H}$, where the term $h\chi_{\{|p|>\alpha\}}$ is evaluated componentwise, *i.e.*,

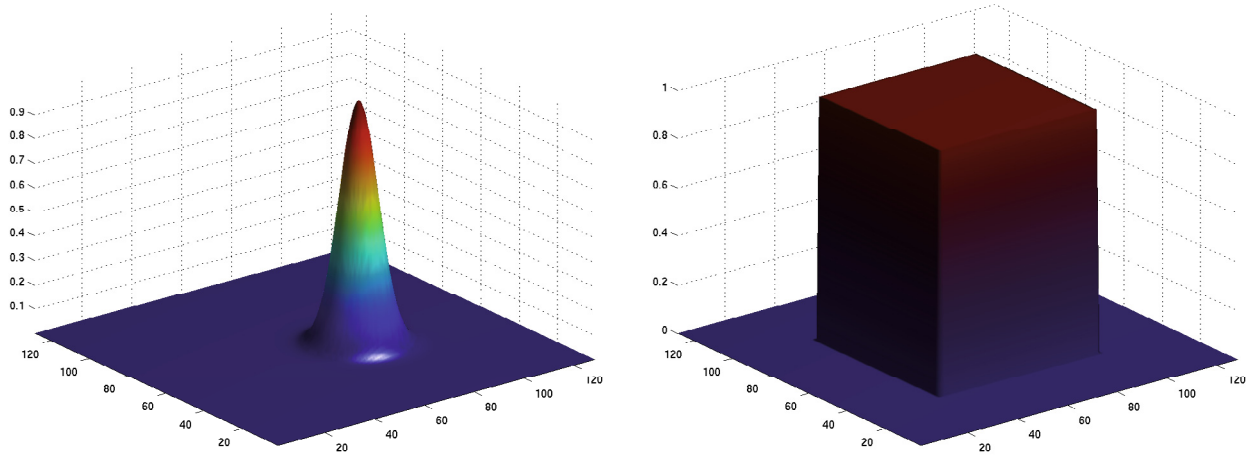
$$(h\chi_{\{|p|>\alpha\}})_i = \begin{cases} h_i(x) & \text{if } |p_i(x)| > \alpha, \\ 0 & \text{if } |p_i(x)| \le \alpha, \end{cases} \tag{3.34}$$

for $i = 1, \dots, n$. One step of the semismooth Newton method consists thus in finding $p^{k+1} \in \mathcal{H}$ such that

$$\left\langle A^* \operatorname{div} p^{k+1}, A^* \operatorname{div} v\right\rangle_{\mathrm{L}^2} + \beta \left\langle \Delta p^{k+1}, \Delta v\right\rangle_{\mathbb{L}^2} + c \left\langle p^{k+1}\chi_{\mathcal{A}_k}, v\right\rangle_{\mathbb{L}^2} = \left\langle z, A^* \operatorname{div} v\right\rangle_{\mathrm{L}^2} + c \left\langle \alpha\chi_{\mathcal{A}^+_k} - \alpha\chi_{\mathcal{A}^-_k}, v\right\rangle_{\mathbb{L}^2} \tag{3.35}$$

holds for all $v \in \mathcal{H}$. The active sets $\mathcal{A}$ and their characteristic functions are defined componentwise as in (3.34). The full Newton method is given as Algorithm 2.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



(A) z_a (B) z_b

FIGURE 1. Test targets.

Since the weak form of the Newton derivative (3.33) by construction defines an inner product on $\mathcal{H}$, the same argument used in Theorem 3.3 yields uniform boundedness of $(\partial_N G)^{-1}$. Hence Algorithm 2 converges locally superlinearly and terminates if the active sets coincide:

Theorem 3.5. *If $\|p_c - p^0\|_{\mathcal{H}}$ is sufficiently small, the iterates p^k of Algorithm 2 converge superlinearly in $\mathcal{H}$ to the solution p_c of (3.31) as $k \to \infty$. Additionally, if $\mathcal{A}^+_{i,k+1} = \mathcal{A}^+_{i,k}$ and $\mathcal{A}^-_{i,k+1} = \mathcal{A}^-_{i,k}$ holds for all $i = 1, \dots, N$, then $p^{k+1} = p_c$.*

Finally, arguing as in Remark 3.2, we obtain that the Moreau-Yosida-regularization of the predual problem is equivalent to adding the penalty term $\frac{1}{2c}\|\nabla u\|^2_{\mathbb{L}^2}$ to the primal problem ($\mathcal{P}_{\mathrm{BV}}$) and minimizing over $u \in \mathrm{H}^1(\Omega)$.

4. Numerical results

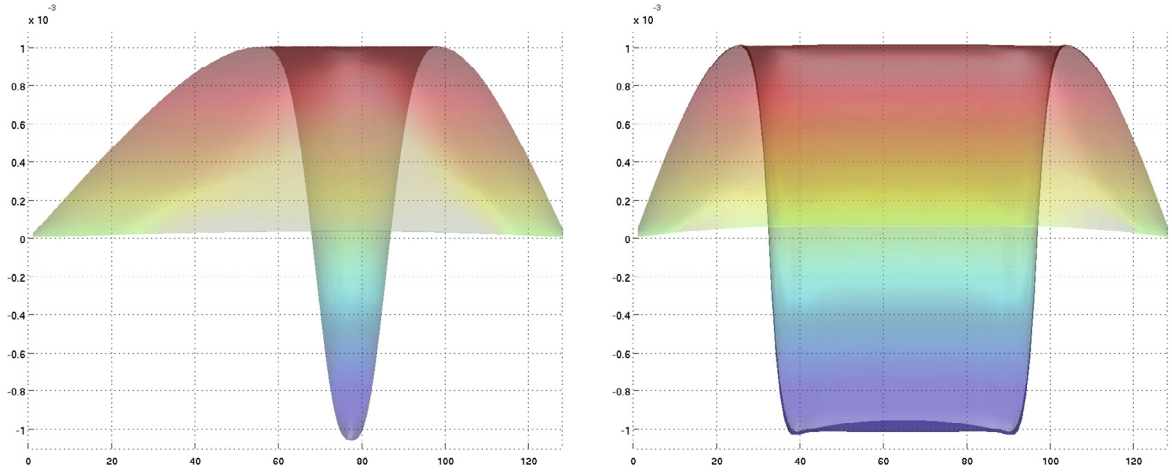
Traditionally, optimal control problems for partial differential equations are formulated with quadratic control costs. When the cost is proportional to the control (or its gradient), it is of interest how this change affects the structure of the optimal controls. The numerical results for a simple model problem presented in this section allow a comparison. Specifically, we consider $A = -\Delta$, the Laplacian with homogeneous Dirichlet conditions on the domain $\Omega = [-1, 1]^2 \subset \mathbb{R}^2$. The differential operators were discretized using standard finite differences on a 128 by 128 grid. To ensure symmetry of the system matrices, the adjoints of A and $-\operatorname{div}$ were taken as the transpose of the corresponding discretization. The implementation was done in Matlab. We consider the following two targets, shown in Figure 1:

$$z_a(x, y) = \mathrm{e}^{-50[(x-0.2)^2+(y+0.1)^2)]}, \tag{4.1}$$

$$z_b(x, y) = \chi_{\{|x|<\frac{1}{2}, |y|<\frac{1}{2}\}}. \tag{4.2}$$

The solutions of the predual problem for these targets will be denoted by p^*_a and p^*_b respectively, and similarly for the resulting optimal controls u^*_a, u^*_b and corresponding states y^*_a, y^*_b.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



(A) p_a^* (B) p_b^*

FIGURE 2. Solutions $p^* \in \mathcal{W}$ of $(\mathcal{P}^*_{\mathcal{M},c})$ for $\alpha = 10^{-3}$, $c = 10^7$. Shown is the projection along x_2.

4.1. Controls in $\mathcal{M}(\Omega)$

We set $\alpha = 10^{-3}$ and $c = 10^7$ and compute the solution of (3.1) using Algorithm 1. The solution of problem $(\mathcal{P}_{\mathcal{M}})$ is then obtained using the first relation of (2.4). The resulting optimal control and corresponding optimal state are shown in Figure 3. The sparsity of the optimal control can be seen (*cf.* Fig. 2, and also Cor. 2.7): the control is zero wherever the dual variable is inactive (*i.e.*, $-\alpha < p^* < \alpha$), negative on the set where the upper bound is active (*i.e.*, $p^* \geq \alpha$), and positive where the lower bound is active (*i.e.*, $p^* \leq -\alpha$). Note that the solution p^* of the regularized problem is allowed to be infeasible, although this can be controlled with larger c.

We compare the controls obtained in $\mathcal{M}(\Omega)$ with the solution of the control problem in $\mathrm{L}^2(\Omega)$:

$$\begin{cases} \min\limits_{u \in \mathrm{L}^2(\Omega)} \dfrac{1}{2} \|y - z\|^2_{\mathrm{L}^2} + \dfrac{\alpha}{2} \|u\|^2_{\mathrm{L}^2} \\ \text{s.t. } Ay = u. \end{cases} \tag{$\mathcal{P}_{\mathrm{L}^2}$}$$

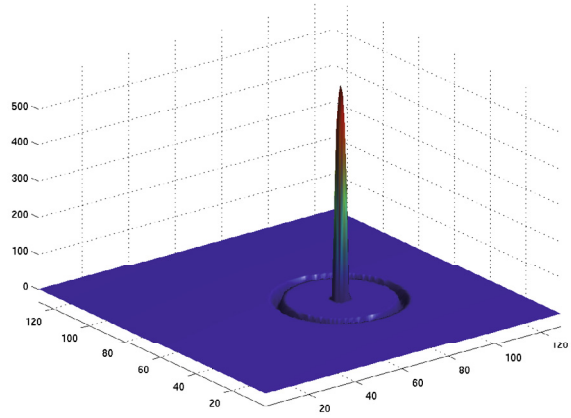
The solution is computed from the optimality system, obtained using standard Lagrangian techniques, which can be written in reduced form as $u + \alpha AA^*u = Az$, if $z \in \mathcal{W}$. The corresponding optimal control and state for the same value of $\alpha = 10^{-3}$ are given in Figure 4. We point out that a better approximation of the target is possible with controls in $\mathcal{M}(\Omega)$ (compare the height of the peaks for target z_a). Note also that the optimal control in $\mathrm{L}^2(\Omega)$ is nonzero almost everywhere, while $u^* \in \mathcal{M}(\Omega)$ is sparse.

We illustrate the superlinear convergence of the Newton method for the case of target z_a with the same parameters as given above. Table 1 gives the error quotients

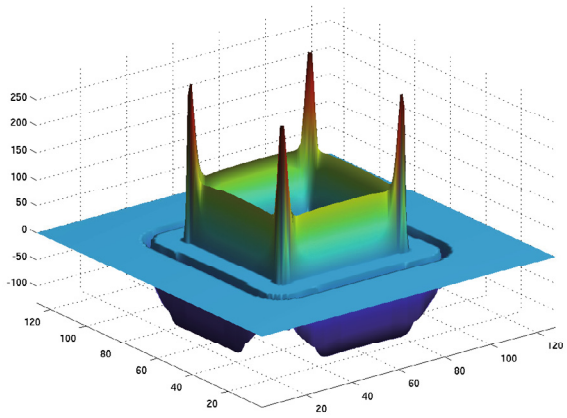
$$e_k := \frac{\left\|p^{k+1} - p^*\right\|_{\mathcal{W}}}{\|p^k - p^*\|_{\mathcal{W}}} \tag{4.3}$$

for the final iterates p^k in the semismooth Newton Algorithm 1. The quotients decrease monotonically, verifying the local superlinear convergence shown in Theorem 3.3.

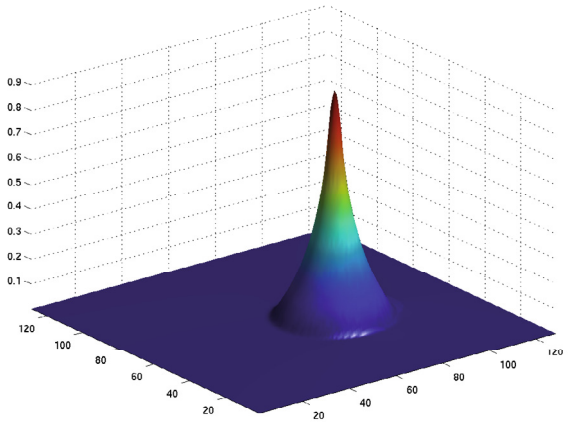
https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



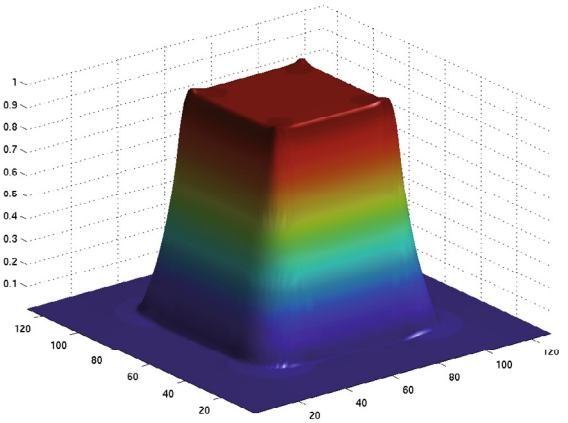
(A) Optimal control $u_a^* \in \mathcal{M}(\Omega)$



(B) Optimal control $u_b^* \in \mathcal{M}(\Omega)$



(C) Optimal state y_a^*



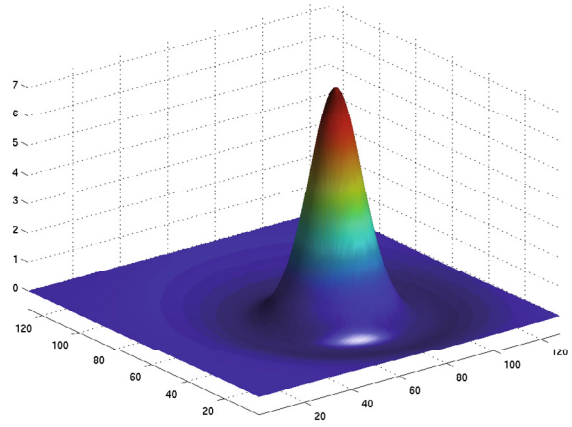
(D) Optimal state y_b^*

FIGURE 3. Solutions of problem ($\mathcal{P}_\mathcal{M}$) calculated *via* predual problem ($\alpha = 10^{-3}$, $c = 10^7$).

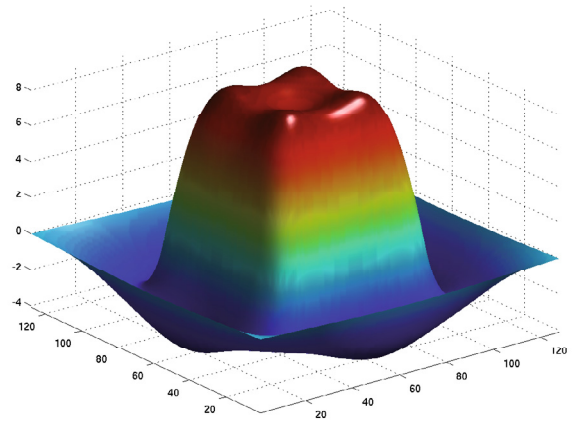
TABLE 1. Convergence of iterates p^k in semismooth Newton method for problem ($\mathcal{P}^*_{\mathcal{M},c}$). Given is the error quotient e_k defined by (4.3) for the final iterates.

k	21	22	23	24	25	26	27
e_k	0.9633	0.6766	0.6164	0.5099	0.3467	0.1048	0

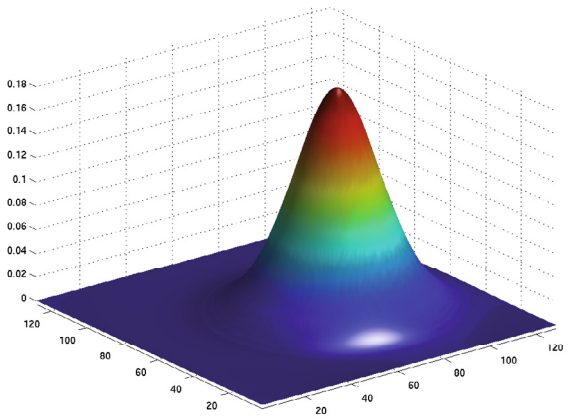
https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



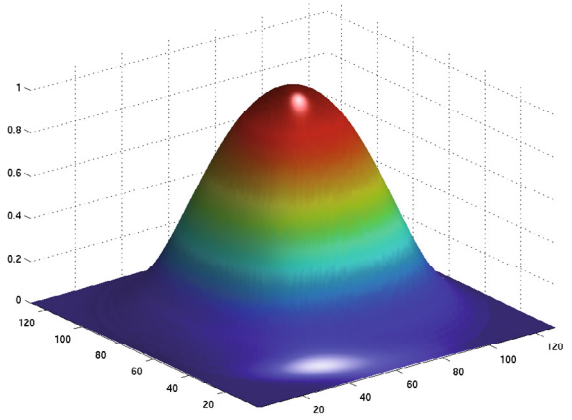
(A) Optimal control $u_a^* \in \mathrm{L}^2(\Omega)$



(B) Optimal control $u_b^* \in \mathrm{L}^2(\Omega)$



(C) Optimal state y_a^*



(D) Optimal state y_b^*

FIGURE 4. Solutions of problem ($\mathcal{P}_{\mathrm{L}^2}$) for $\alpha = 10^{-3}$.

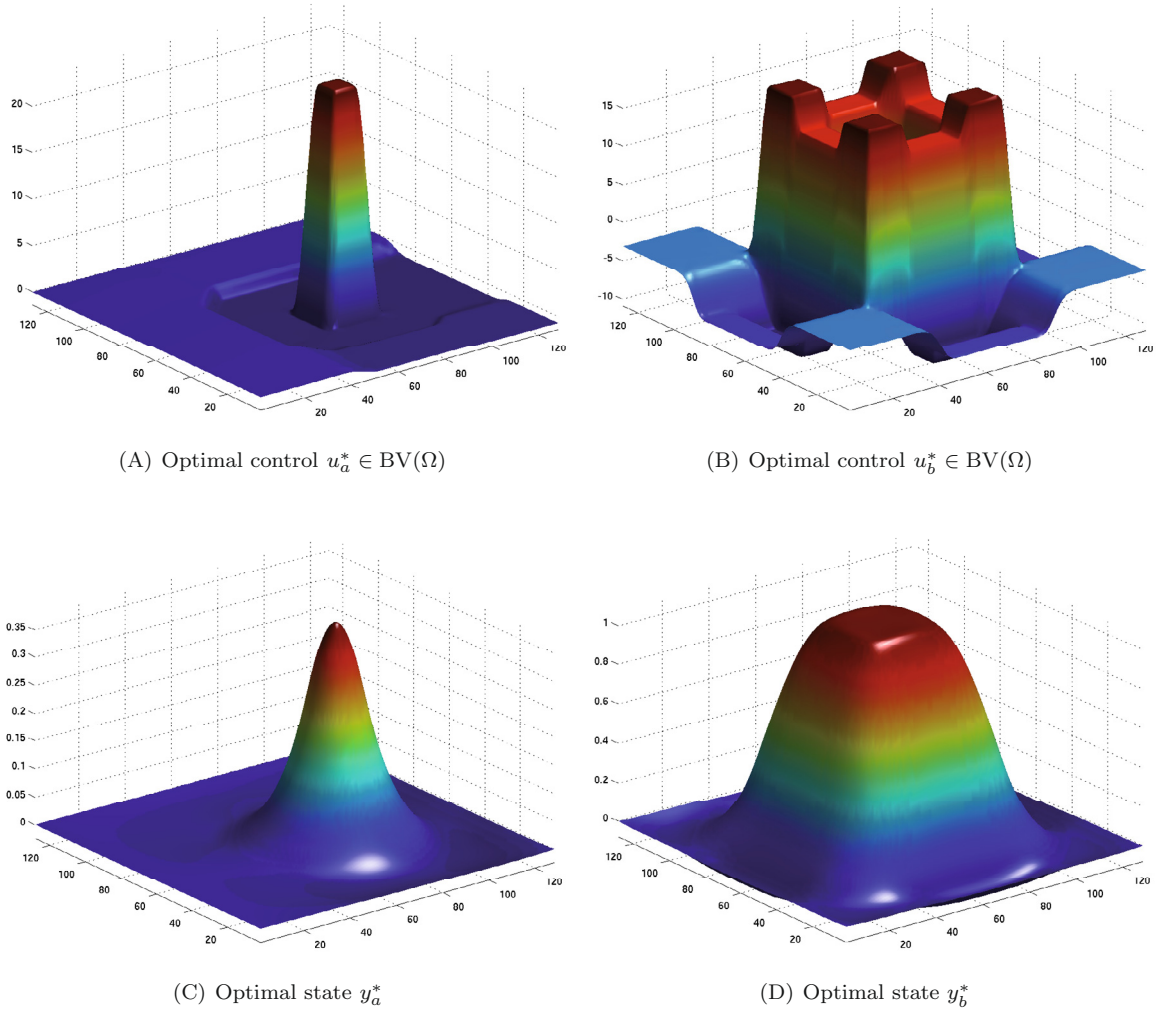
4.2. Controls in BV(Ω)

We repeat the computations for problem ($\mathcal{P}^*_{\mathrm{BV},c}$), setting $\alpha = 10^{-4}$, $\beta = 10^{-1}$ and $c = 10^7$. Here, we compare with the optimal control in $\mathrm{H}^1_0(\Omega)$:

$$\begin{cases} \min\limits_{u \in \mathrm{H}^1_0(\Omega)} \dfrac{1}{2} \|y - z\|^2_{\mathrm{L}^2} + \dfrac{\alpha}{2} \|u\|^2_{\mathrm{H}^1(\Omega)} \\ \text{s.t. } Ay = u, \end{cases} \tag{$\mathcal{P}_{\mathrm{H}^1}$}$$

where the solution is computed from $u - \alpha AA^* \Delta u = Az$, if $z \in \mathcal{W}$. The resulting controls in $\mathrm{BV}(\Omega)$ and $\mathrm{H}^1_0(\Omega)$ and corresponding states are given in Figures 5 and 6, respectively. It can be clearly seen that the $\mathrm{BV}(\Omega)$ cost favors piecewise constant controls; this phenomenon is well-known in the mathematical imaging community as staircasing (*cf.* [10]). Again, the target is better attained by controls in $\mathrm{BV}(\Omega)$ compared with $\mathrm{H}^1_0(\Omega)$.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



(A) Optimal control $u_a^* \in \mathrm{BV}(\Omega)$

(B) Optimal control $u_b^* \in \mathrm{BV}(\Omega)$

(C) Optimal state y_a^*

(D) Optimal state y_b^*

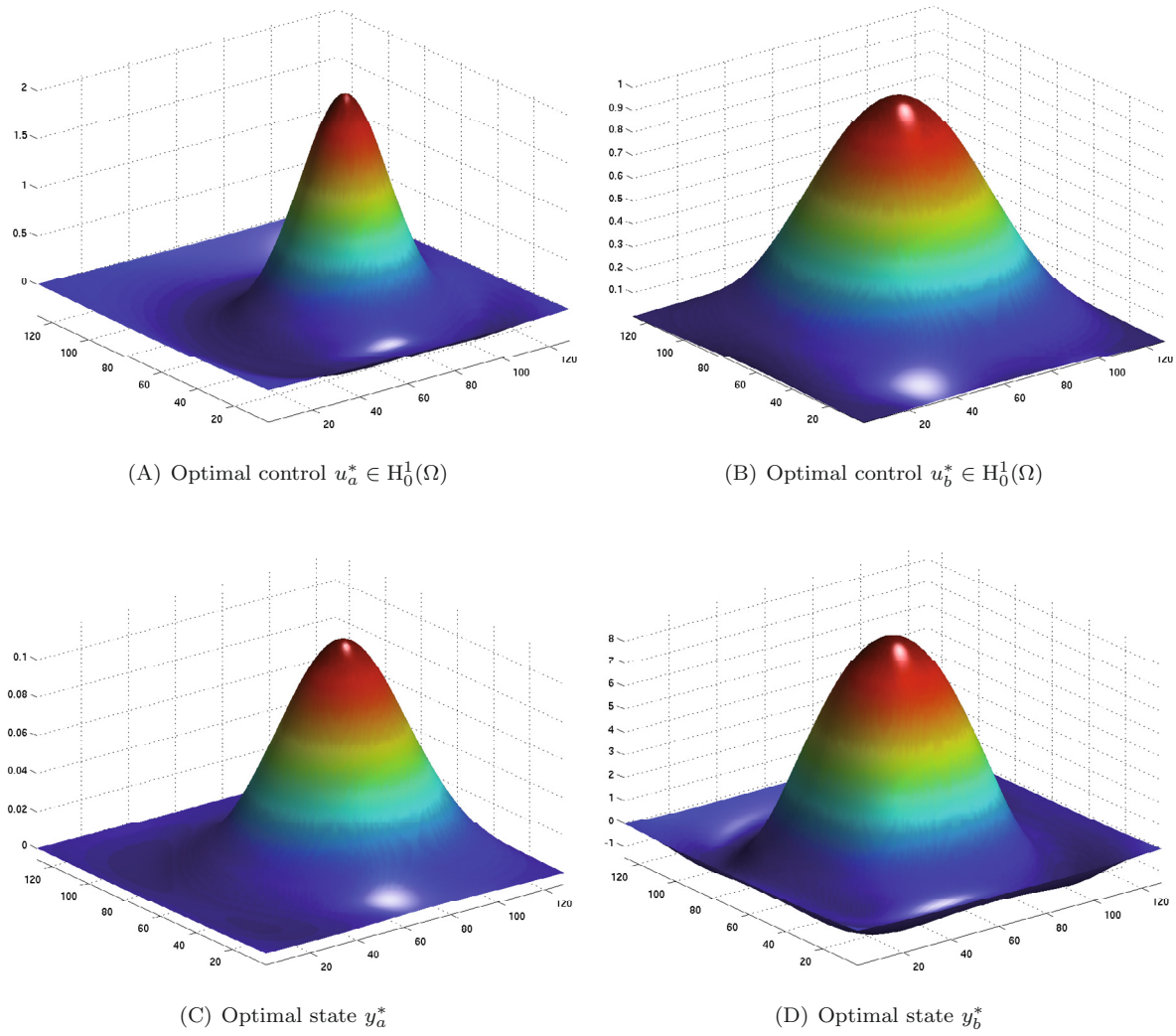
FIGURE 5. Solutions of problem ($\mathcal{P}_{\mathrm{BV}}$) calculated *via* predual problem ($\alpha = 10^{-4}$, $\beta = 10^{-1}$, $c = 10^7$).

To show the effect of the control cost in the case of $\mathrm{BV}(\Omega)$, we increase α from 10^{-4} to 10^{-3} for the target z_b. The optimal control and corresponding state are shown in Figure 7. The control is now constant over a much larger region, at the cost of being further from the target. For this example, we can easily illustrate Corollary 2.14: The derivative with respect to x_1 or x_2 of the control u^* is zero in regions where the box constraint for the corresponding component of the predual solution $p^* = (p_1^*, p_2^*)$ is inactive (*cf.* Fig. 8).

5. Conclusion

We have presented a framework for the efficient solution of elliptic optimal control problems in non-reflexive Banach spaces such as the space of bounded Radon measures or of functions of bounded variation. Specifically,

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



(A) Optimal control $u_a^* \in \mathrm{H}_0^1(\Omega)$

(B) Optimal control $u_b^* \in \mathrm{H}_0^1(\Omega)$

(C) Optimal state y_a^*

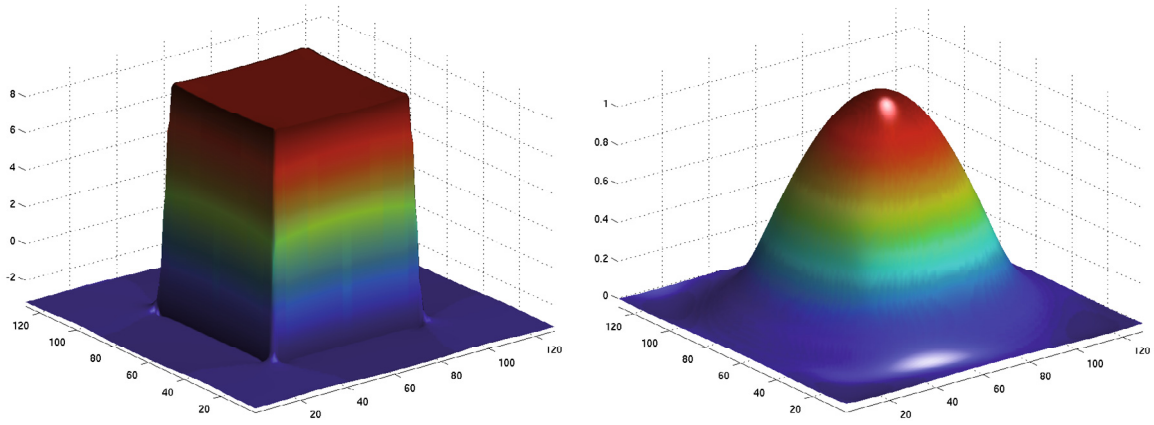
(D) Optimal state y_b^*

FIGURE 6. Solutions of problem ($\mathcal{P}_{\mathrm{H}^1}$) for $\alpha = 10^{-4}$.

the Fenchel duality theorem allows the reformulation of these (non-differentiable) problems as smooth box constrained problems in a Hilbert space. The corresponding optimality systems can be solved with a semismooth Newton method, which converges superlinearly after regularization. We also demonstrated the structural differences between the optimal controls for these non-smooth problems and the solutions of the corresponding quadratic formulations (*i.e.*, $\mathrm{L}^2(\Omega)$ and $\mathrm{H}_0^1(\Omega)$).

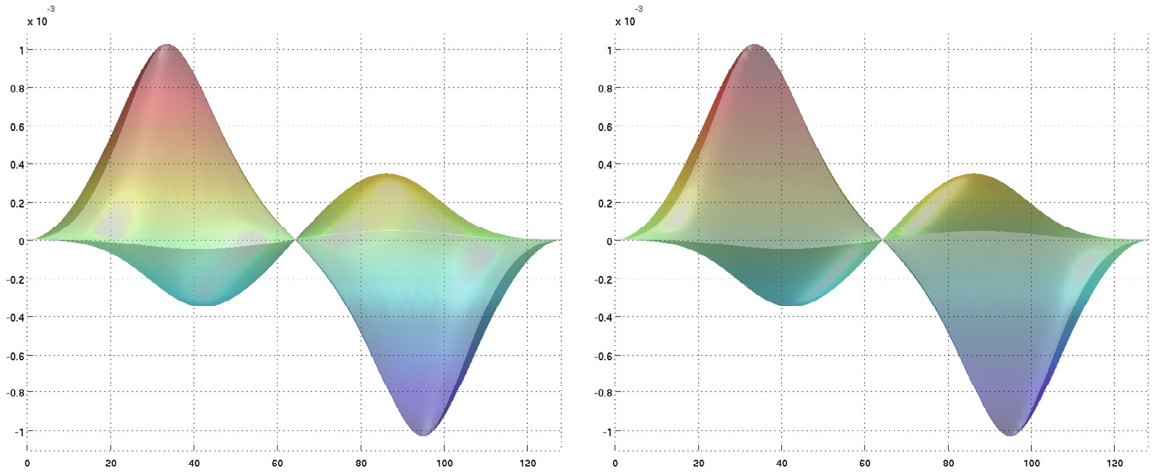
The proposed approach can be extended to time-dependent (*e.g.*, parabolic) problems as well, which will be investigated in a subsequent work.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press



(A) Optimal control $u_b^* \in \mathrm{BV}(\Omega)$ (B) Optimal state y_b^*

FIGURE 7. Solution of problem ($\mathcal{P}_{\mathrm{BV}}$) for target z_b with $\alpha = 10^{-3}$.



(A) Projection of p_1^* along x_2 (B) Projection of p_2^* along x_1

FIGURE 8. Solution $p_b^* = (p_1^*, p_2^*) \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ of ($\mathcal{P}^*_{\mathrm{BV},c}$) with $\alpha = 10^{-3}$, $c = 10^7$.

A. CONVERGENCE OF MOREAU-YOSIDA-REGULARIZATION OF ($\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2}$)

In this appendix we consider the regularization of ($\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2}$):

$$\min_{p\in\mathcal{H}} \frac{1}{2}\left\|A^*\operatorname{div} p + z\right\|^2_{\mathrm{L}^2} + \frac{\beta}{2}\left\|p\right\|^2_{\mathbb{L}^2} + \frac{1}{2c}\left\|p\right\|^2_{\mathcal{W}^n} - \frac{1}{2}\left\|z\right\|^2_{\mathrm{L}^2} + \frac{1}{2c}\left\|\max(0, c(p-\alpha))\right\|^2_{\mathbb{L}^2} + \frac{1}{2c}\left\|\min(0, c(p+\alpha))\right\|^2_{\mathbb{L}^2}, \quad (\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2,c})$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

$\beta > 0$ fixed. For the sake of presentation, we have used the same parameter for the Moreau-Yosida regularization and the $\mathcal{W}^n$-smoothing term, but the same result holds if we replace $\frac{1}{2c}\|p\|^2_{\mathcal{W}^n}$ by $\frac{\gamma}{2}\|p\|^2_{\mathcal{W}^n}$ and take the limit as $c \to \infty$, $\gamma \to 0$.

The corresponding optimality system is

$$\begin{cases} \langle A^* \operatorname{div} p_c, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \beta \langle p_c, v\rangle_{\mathbb{L}^2} + \dfrac{1}{c}\langle \Delta p_c, \Delta v\rangle_{\mathbb{L}^2} - \langle z, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \langle \lambda_c, v\rangle_{\mathcal{H}^*,\mathcal{H}} = 0, \\ \lambda_c = \max(0, c(p_c - \alpha)) + \min(0, c(p_c + \alpha)), \end{cases} \tag{A.1}$$

for all $v \in \mathcal{H}$. This equation has a unique solution $(\lambda_c, p_c) \in \mathcal{H}^* \times \mathcal{H}$. We wish to show convergence as $c \to \infty$ to the unique solution $(\lambda^*, p^*) \in \mathrm{H}^2_{\mathrm{div}}(\Omega)^* \times \mathrm{H}^2_{\mathrm{div}}(\Omega)$ of the optimality system for $(\mathcal{P}^*_{\mathrm{BV},\mathrm{L}^2})$:

$$\begin{cases} \langle A^* \operatorname{div} p^*, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \beta \langle p^*, v\rangle_{\mathbb{L}^2} - \langle z, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \langle \lambda^*, v\rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*,\mathrm{H}^2_{\mathrm{div}}} = 0, \\ \langle \lambda^*, p^* - p\rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*,\mathrm{H}^2_{\mathrm{div}}} \le 0, \end{cases} \tag{A.2}$$

for all $p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$ with $\|p\|_{\mathbb{L}^\infty} \le \alpha$.

The proof is similar to that of Theorem 3.1. Arguing as above, we have that

$$\langle \lambda_c, p_c\rangle_{\mathrm{L}^2} \ge \frac{1}{c}\|\lambda_c\|^2_{\mathrm{L}^2} \tag{A.3}$$

and hence from (A.1) that

$$\|A^* \operatorname{div} p_c\|^2_{\mathrm{L}^2} + \frac{1}{c}\|p_c\|^2_{\mathcal{W}^n} + \beta\|p_c\|^2_{\mathbb{L}^2} + \frac{1}{c}\|\lambda_c\|^2_{\mathrm{L}^2} \le \|A^* \operatorname{div} p_c\|_{\mathrm{L}^2}\|z\|_{\mathrm{L}^2}. \tag{A.4}$$

This implies that the $\mathrm{H}^2_{\mathrm{div}}(\Omega)$-norm of p_c is bounded uniformly in c by $\|z\|_{\mathrm{L}^2}$, and that

$$\|\lambda_c\|_{\mathcal{H}^*} \le 2 \sup_{\substack{v\in\mathcal{H},\\ \|v\|_{\mathcal{H}}\le 1}} \|A^* \operatorname{div} v\|_{\mathrm{L}^2}\|z\|_{\mathrm{L}^2} < \infty. \tag{A.5}$$

It follows that (λ_c, p_c) converges weakly subsequentially in $\mathcal{H}^* \times \mathrm{H}^2_{\mathrm{div}}(\Omega)$ to a $(\tilde\lambda, \tilde p) \in \mathcal{H}^* \times \mathrm{H}^2_{\mathrm{div}}(\Omega)$, which satisfies

$$\langle A^* \operatorname{div} \tilde p, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \beta\langle \tilde p, v\rangle_{\mathbb{L}^2} - \langle z, A^* \operatorname{div} v\rangle_{\mathrm{L}^2} + \left\langle \tilde\lambda, v\right\rangle_{\mathcal{H}^*,\mathcal{H}} = 0 \tag{A.6}$$

for all $v \in \mathcal{H}$. By the density of $\mathcal{H}$ in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$ (*via* $(\mathrm{C}^\infty_0(\Omega))^n \subset \mathcal{H}$), equation (A.6) holds for all $v \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$, and we can identify $\tilde\lambda$ with an element in $\mathrm{H}^2_{\mathrm{div}}(\Omega)^*$ (replacing the duality pairing with the one in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$).

The feasibility of $\tilde p$ follows exactly as in the proof of Theorem 3.1. Similarly, we deduce from the optimality of p_c and the density of $\mathcal{H}$ in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$ that

$$\limsup_{c\to\infty} \frac{1}{2}\|A^* \operatorname{div} p_c + z\|^2_{\mathrm{L}^2} + \frac{\beta}{2}\|p_c\|^2_{\mathbb{L}^2} \le \frac{1}{2}\|A^* \operatorname{div} \tilde p + z\|^2_{\mathrm{L}^2} + \frac{\beta}{2}\|\tilde p\|^2_{\mathbb{L}^2} \tag{A.7}$$

holds. The convergence of p_c to $\tilde p$ is therefore strong in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$, and we can pass to the limit in

$$\langle \lambda_c, p - p_c\rangle_{\mathcal{H}^*,\mathcal{H}} \le 0 \quad \text{for all feasible } p \in \mathcal{H}, \tag{A.8}$$

to obtain

$$\left\langle \tilde\lambda, p - \tilde p\right\rangle_{\mathcal{H}^*,\mathcal{H}} \le 0 \quad \text{for all feasible } p \in \mathcal{H}. \tag{A.9}$$

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press

Again, the density of $\mathcal{H}$ in $\mathrm{H}^2_{\mathrm{div}}(\Omega)$ and the fact that $\tilde{\lambda} \in \mathrm{H}^2_{\mathrm{div}}(\Omega)^*$ allows us to conclude that

$$\left\langle \tilde{\lambda}, p - \tilde{p} \right\rangle_{\mathrm{H}^2_{\mathrm{div}}{}^*, \mathrm{H}^2_{\mathrm{div}}} \leq 0 \tag{A.10}$$

holds for all feasible $p \in \mathrm{H}^2_{\mathrm{div}}(\Omega)$. Therefore, $(\tilde{\lambda}, \tilde{p}) \in \mathrm{H}^2_{\mathrm{div}}(\Omega)^* \times \mathrm{H}^2_{\mathrm{div}}(\Omega)$ satisfies (A.2), and $\tilde{p} = p^*$ and $\tilde{\lambda} = \lambda^*$ follows from the uniqueness of its solution.

References

[1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, *Oxford Mathematical Monographs*. Oxford University Press, New York, USA (2000).

[2] C. Amrouche, P.G. Ciarlet and P. Ciarlet, Jr., Vector and scalar potentials, Poincaré's theorem and Korn's inequality. *C. R. Math. Acad. Sci. Paris* **345** (2007) 603–608.

[3] H. Attouch, G. Buttazzo and G. Michaille, *Variational analysis in Sobolev and BV spaces*, *MPS/SIAM Series on Optimization* **6**. Society for Industrial and Applied Mathematics, Philadelphia, USA (2006).

[4] H. Brezis, *Analyse fonctionnelle*, *Collection Mathématiques Appliquées pour la Maîtrise*. Masson, Paris, France (1983).

[5] G. Chavent and K. Kunisch, Regularization of linear least squares problems by total bounded variation. *ESAIM: COCV* **2** (1997) 359–376.

[6] I. Ekeland and R. Témam, *Convex analysis and variational problems*. Society for Industrial and Applied Mathematics, Philadelphia, USA (1999).

[7] M. Hintermüller and G. Stadler, An infeasible primal-dual algorithm for total bounded variation-based inf-convolution-type image restoration. *SIAM J. Sci. Comput.* **28** (2006) 1–23.

[8] M. Hintermüller, K. Ito and K. Kunisch, The primal-dual active set strategy as a semismooth Newton method. *SIAM J. Optim.* **13** (2002) 865–888.

[9] K. Ito and K. Kunisch, *Lagrange multiplier approach to variational problems and applications*, *Advances in Design and Control* **15**. Society for Industrial and Applied Mathematics, Philadelphia, USA (2008).

[10] W. Ring, Structural properties of solutions to total variation regularization problems. *ESAIM: M2AN* **34** (2000) 799–810.

[11] G. Stadler, Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices. *Comp. Optim. Appl.* **44** (2009) 159–181.

[12] G. Stampacchia, Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus. *Ann. Inst. Fourier (Grenoble)* **15** (1965) 189–258.

[13] R. Témam, *Navier-Stokes equations*. AMS Chelsea Publishing, Providence, USA (2001).

[14] M. Ulbrich, Semismooth Newton methods for operator equations in function spaces. *SIAM J. Optim.* **13** (2002) 805–842.

[15] G. Vossen and H. Maurer, On L^1-minimization in optimal control and applications to robotics. *Optimal Control Appl. Methods* **27** (2006) 301–321.

https://doi.org/10.1051/cocv/2010003 Published online by Cambridge University Press